Engineering *Escherichia coli* for the Novel and Enhanced Biosynthesis of Phenol, Catechol, and Muconic Acid

by

Brian Thompson

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2017 by the
Graduate Supervisory Committee:

David R. Nielsen, Chair
Brent Nannenga
Matthew Green
Xuan Wang
Tae Seok Moon

ARIZONA STATE UNIVERSITY

AUGUST 2017

ABSTRACT

The engineering of microbial cell factories capable of synthesizing industrially relevant chemical building blocks is an attractive alternative to conventional petrochemical-based production methods. This work focuses on the novel and enhanced biosynthesis of phenol, catechol, and muconic acid (MA). Although the complete biosynthesis from glucose has been previously demonstrated for all three compounds, established production routes suffer from notable inherent limitations. Here, multiple pathways to the same three products were engineered, each incorporating unique enzyme chemistries and/or stemming from different endogenous precursors. In the case of phenol, two novel pathways were constructed and comparatively evaluated, with titers reaching as high as 377 ± 14 mg/L at a glucose yield of 35.7 ± 0.8 mg/g. In the case of catechol, three novel pathways were engineered with titers reaching 100 ± 2 mg/L. Finally, in the case of MA, four novel pathways were engineered with maximal titers reaching 819 ± 44 mg/L at a glucose yield of 40.9 ± 2.2 mg/g. Furthermore, the unique flexibility with respect to engineering multiple pathways to the same product arises in part because these compounds are common intermediates in aromatic degradation pathways. Expanding on the novel pathway engineering efforts, a synthetic 'metabolic funnel' was subsequently constructed for phenol and MA, wherein multiple pathways were expressed in parallel to maximize carbon flux toward the final product. Using this novel 'funneling' strategy, maximal phenol and MA titers exceeding 0.5 and 3 g/L, respectively, were achieved, representing the highest achievable production metrics products reported to date.

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Dr. David Nielsen, for his continued support and guidance throughout my graduate school career. I am extremely appreciative of his technical insight, patience, and calming voice of reason during the inevitable ups and downs that accompany research endeavors. This work would not have been possible without his support, and I will forever be grateful of his unwavering dedication. I would also like to thank Drs. Brent Nannenga, Matt Green, Xuan Wang, and Tae Seok Moon for taking time out of their busy schedules to serve on my committee and provide useful guidance and insight. Finally, I would like to thank all my friends and lab colleagues who helped make this experience fun and exciting. A special thanks to Anirudh Vasudevan, Michael Machas, Kyle Staggs, Danny Herschel, Bohan Shan, Karthik Pushpavanam, Zach Dookeran, Chris Gregson, Marwan Osman, Vick Syradi, and Drs. Matt Christensen, Sai Pavan, Sudhakar Godeshala, Chris Jones – from racquetball games to encouraging my coffee addiction, I could not have done this without your friendship and support.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1. CREATING PATHWAYS TOWARDS AROMATIC BUILDING BLOCKS AND FINE CHEMICALS 1

    1.1. Introduction 2

    1.2. A Diverse Precursor Pool Supports a Range of Aromatic Products 4

    1.3. Constructing Aromatic Biosynthesis Pathways from Robust Enzyme 'Parts' 5

    1.4. Exploiting Modularity to Enhance and Expand Aromatic Chemical Production 7

    1.5. Multiple Pathways, One Product 9

    1.6. Inherent Challenges, Future Outlooks, and Conclusions 11

    1.7. Dissertation Organization 12

2. ENGINEERING AND COMPARISON OF NON-NATURAL PATHWAYS FOR MICROBIAL PHENOL PRODUCTION 13

    2.1. Introduction 14

    2.2. Materials and Methods 18

        2.2.1. Bacterial Strains and Media 18

        2.2.2. Plasmid Construction 19

CHAPTER Page

2.2.3. Strain Construction 20
2.2.4. Assaying Candidate Enzyme Activity Using Whole Resting Cells ... 22
2.2.5. Biosynthesis of Phenol from Glucose in Shake Flask Cultures 22
2.2.6. HPLC Analysis 23
2.2.7. Elementary Flux Mode Analysis 23
2.3. Results and Discussion 24
2.3.1. Identifying and Characterizing the Requisite Pathway Enzymes ... 24
2.3.2. Phenol Pathway Construction in Wild-type E. coli 29
2.3.3. Increasing Intracellular Availability of Precursor Chorismate in Support of Phenol Production 30
2.3.4. Further Strain Engineering in Support of Phenol Biosynthesis 34
2.3.5. Comparing Relative Performance of Established and Novel Phenol Biosynthesis Pathways 36
2.4. Conclusions 39
3. ENGINEERING A NETWROK OF PATHWAYS FOR THE NOVEL BIOSYNTHESIS OF MUCONIC ACID 41
3.1. Introduction 42
3.2. Materials and methods 44
3.2.1. Strains and Media 44
3.2.2. Plasmid Construction 45
3.2.3. Strain Construction 46

CHAPTER Page

3.2.4. Thermodynamic and Elementary Flux Mode Analysis 46

3.2.5. Assaying Phenol Hydroxylase and Catechol 1,2-Dioxygenase Activity Using Whole Resting Cells 47

3.2.6. Catechol and MA Production from Glucose 49

3.2.7. HPLC Metabolite Analysis 50

3.3. Results and Discussion 51

3.3.1. Novel Pathway Identification and Theoretical Comparison 51

3.3.2. Screening for and Characterizing Phenol Hydroxylase Activity in E. coli 56

3.3.3. Investigating Phenol-Dependent Catechol Production in E. coli 59

3.3.4. Investigating MA Production via Newly Engineered Pathways 61

3.3.5. Conclusions 63

4. EXPLORING A SYNTHETIC METABOLIC FUNNELING APPROACH TO ENHANCE PHENOL AND MUCONIC ACID BIOSYNTHESIS 64

4.1. Introduction 65

4.2. Materials and Methods 67

4.2.1. Strains and Media 67

4.2.2. Plasmid Construction 70

4.2.3. Strain Construction 71

4.2.4. Phenol and MA Production 71

4.2.5. HPLC Metabolite Analysis 72

CHAPTER Page

4.3. Results and Discussion 73

4.3.1. Demonstrating Phenol and MA Production via Synthetic 'Metabolic Funneling' 73

4.3.2. Host Engineering to Enhance Precursor Availability 76

4.3.3. Synthetic 'Funneling' for Renewable Chemical Production 79

4.4. Conclusions 81

5. DISCUSSION AND FUTURE WORK 82

5.1. Engineering a Dynamic Phenotype Switch 83

5.1.1. Transcriptional Control 87

5.1.2. Translational Control 91

5.1.3. Post-translational Control 92

5.1.4. Phenotype Switch Design Considerations 93

5.1.5. Phenotype Switch Characterization 96

5.1.6. Anticipated Outcomes 97

5.1.7. Pitfalls and Proposed Solutions 98

5.2. Dynamic Flux Regulation for Phenol and MA Production 99

5.3. Conclusions 101

REFERENCES 102

APPENDIX Page

A SUPPLEMENTAL INFORMATION TO ENGINEERING AND COMPARISON OF NON-NATURAL PATHWAYS FOR MICROBIAL PHENOL PRODUCTION 114

B SUPPLEMENTAL INFORMATION TO ENGINEERING NOVEL PATHWAYS FOR THE BIOSYNTHESIS OF MUCONIC ACID 121

C SUPPLEMENTAL INFORMATION TO EXPLORING A SYNTHETIC METABOLIC FUNNELING APPROACH TO ENHANCE PHENOL AND MUCONIC ACID BIOSYNTHESIS 131

D PRELIMINARY DATA SUPPORTING THE ENGINEERING OF A DYNAMIC PHENOTYPE CONTROL SWITCH 133

LIST OF TABLES

Table Page

2.1 Strains and Plasmids for Phenol Production 21

3.1 Strains and Plasmids Used for MA Production 48

3.2 Strains Engineered for Catechol and MA Biosynthesis 49

3.3 Catechol Production via Proposed Novel Pathways 61

3.4 Novel Muconic Acid Production 63

4.1 Strains and Plasmids Used for Phenol and MA Production 69

4.2 Strains Engineered for Phenol and MA Metabolic 'Funneling' 70

4.3 Evaluating Metabolic 'Funneling' for Phenol and MA Biosynthesis 75

4.4 Comparing Engineering Strategies for Phenol and MA Production 77

4.5 3-hydroxypropionate Biosynthesis Pathways 80

5.1 Dynamic Phenotype Control Mechanisms 86

LIST OF FIGURES

Figure Page

1.1 Metabolic Network of Aromatic Biosynthesis Pathways. 3

1.2 Chorismate Derived Aromatic Biosynthesis. 6

1.3 Phenol Biosynthesis Pathways. 10

2.1 Engineered Phenol Biosynthesis. 16

2.2 Gibbs Free Energy. 25

2.3 Recombinant Enzyme Activity 28

2.4 Wild-type Phenol Production 30

2.5 Host Strain Engineering to Support Phenol Production. 34

2.6 Alternative Host Selection for Phenol Production. 35

3.1 Five Non-natural Pathways Engineered for Muconic Acid Biosynthesis 54

3.2 Comparing the Change in Gibbs Free Energy of Reaction 55

3.3 Screening Candidate Phenol Hydroxylase Activity 57

4.1 Engineering a Synthetic Metabolic 'Funnel' 67

4.2 Engineering a Synthetic Funnel for Phenol Production 74

5.1 Metabolic Flux Distribution for Enhanced Product Biosynthesis 84

5.2 Dynamic Flux Control Response 85

5.3 Synthetic CRISPRi 88

5.4 Engineered Toggle Switch 89

5.5 Engineering Dynamic Homologous Recombination 90

5.6 Small Synthetic Regulatory RNA (sRNA) 92

5.7 Tunable Protein Degradation 93

5.8 Dynamic Phenotype Switch Output 98

1. CREATING PATHWAYS TOWARDS AROMATIC BUILDING BLOCKS AND FINE CHEMICALS

Abstract

Aromatic compounds represent a broad class of chemicals with a range of industrial applications, all of which are conventionally derived from petroleum feedstocks. However, owing to a diversity of available pathway precursors along with natural and engineered enzyme 'parts', microbial cell factories can be engineered to create alternative, renewable routes to many of the same aromatic products. Drawing from the latest tools and strategies in metabolic engineering and synthetic biology, such efforts are becoming an increasingly systematic practice, while continued efforts promise to open new doors to an ever-expanding range and diversity of renewable chemical and material products. This chapter will highlight recent and notable achievements regarding the microbial production of aromatic chemicals and provide an organizational outline of this dissertation.

This chapter was published as:

Thompson, B., Machas, M., Nielsen, D.R. Creating Pathways Towards Aromatic Building Blocks and Fine Chemicals (2015). *Curr Opin Biotechnol*, 36:1-7.

1.1. Introduction

As the focus of many researchers in the biorenewables sector continues to transition from biofuels to biochemicals, the motivation behind the shift is clear: bulk and fine chemicals are of higher value than fuels (per unit mass) and serve significantly smaller markets (based on total carbon demand). Thus, although the total chemicals market utilizes only 5% as much carbon, the net economic value nearly equals that of fuels (Nikolau et al. 2008). As a result, aggressive growth in this sector has been experienced worldwide in recent years – a trend that is expected to continue. In the United States, for example, production targets of bio-based chemicals and materials have been set at 18 and 25% of total production levels by 2020 and 2030, respectively (Perlack et al. 2005). In Europe, meanwhile, net chemical production from biomass is expected to reach 30% by 2030 (up from 10% in 2013) (Council 2006). In addition to the continued development of established bioproducts, further growth and diversification amongst the renewable biochemicals market will occur most significantly through the engineering of novel microbial biocatalysts capable of synthesizing new and useful products as renewable replacements to conventional petrochemicals. Several recent examples illustrate how this approach has already been used to broaden the spectrum of chemical products that can synthesized from renewable biomass, including 1,4-butanediol (Yim et al. 2011), iso-butanol (Atsumi et al. 2010), isoprene (Zhao et al. 2011), and various organic acids (Adkins et al. 2013; Shiue and Prather 2014; Zhang et al. 2011).

Aromatic compounds represent another important class of conventional petrochemicals that have recently emerged as promising targets for metabolic pathway engineering. Conventionally derived from non-renewable petroleum feedstocks including benzene, toluene, and xylenes (BTX), aromatic chemicals serve a vast market and find extensive applications in the production of various solvents,

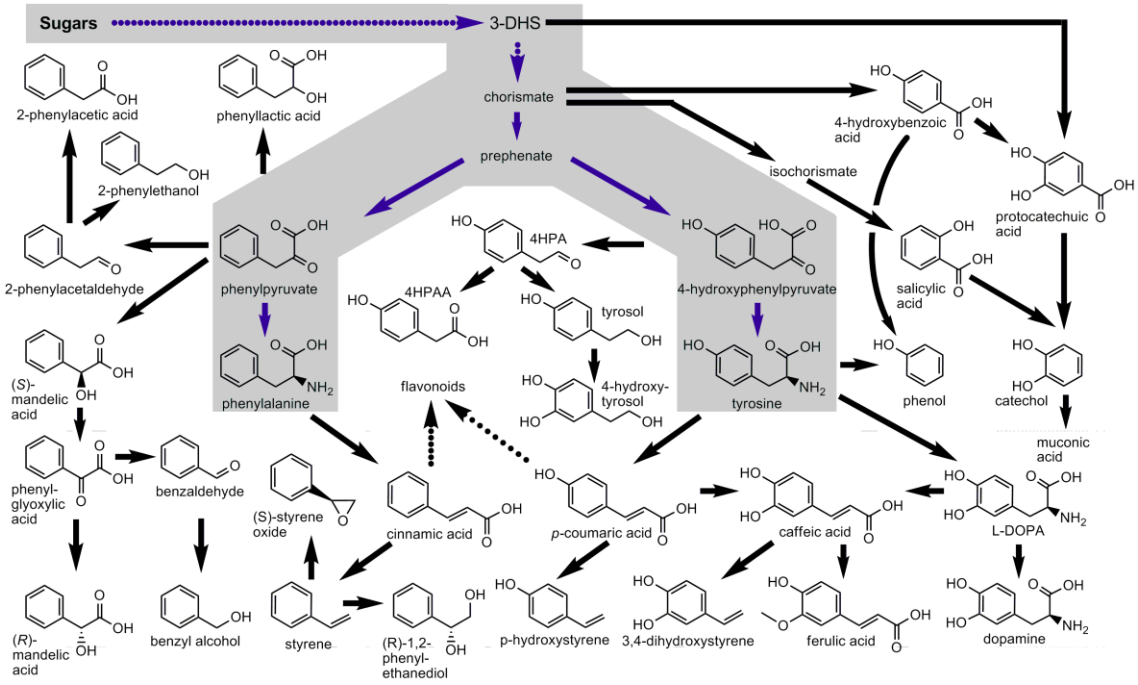


Figure 1.1 Metabolic Network of Aromatic Biosynthesis Pathways.By applying metabolic engineering and synthetic biology tools, researchers have engineered a collective 'network' of biosynthetic pathways to enable the microbial biosynthesis of an array of useful aromatic chemicals from renewable, biomass-derived sugars. Illustrated are both functionally reconstructed, naturally occurring pathways as well as novel, non-natural pathways. Shaded regions depict endogenous metabolism. Note: Trp derived pathways are excluded for clarity and space.

plastics, and fine chemicals (Myers 2007). In 2010, for example, global annual consumption of benzene alone was estimated at over 40 million tons and growing steadily (Nexant 2011). Using microbes such as *Escherichia coli* and *Saccharomyces cerevisiae* as hosts, novel enzyme pathways are alternatively being engineered to enable the novel biosynthesis of an increasing array of aromatic chemicals directly from renewable, biomass-derived feedstocks (Figure 1.1). In general, the novel ability of microbes to produce aromatic biochemicals has been accomplished through either *a)* the functional reconstruction of natural-occurring pathways (typically from higher and/or less tractable organism; e.g., plants and fungi) or *b) de novo* pathway engineering. Supported by advancements in bioinformatics, systems and synthetic

biology, protein engineering, and 'omics' tools, the latter approach continues to emerge as an increasingly systematic practice, and one that is uniquely enabling the biosynthesis of non-natural aromatic chemicals. This short review aims to provide an overview of recent and impactful efforts to this end, with particular emphasis given to enabling factors uniquely associated with the biosynthesis of aromatic chemicals.

1.2. A Diverse Precursor Pool Supports a Range of Aromatic Products

With few exceptions (perhaps only phloroglucinol, a phenolic compound derived directly from malonyl-CoA (Achkar et al. 2005)), aromatic pathways have been constructed by utilizing endogenous precursors derived from the ubiquitous shikimate pathway. Accordingly, achieving appreciable production of any aromatic target first requires metabolite flux to be enhanced through this otherwise tightly regulated pathway: a prerequisite condition that has been met in numerous microbial platforms, typically as a result of both combinatorial (e.g., metabolic evolution) and rational approaches (Juminaga et al. 2012; McKenna et al. 2014; Santos et al. 2012). A key and relatively unique factor enabling access to a broad range of established and putative final aromatic products is in fact the diverse pool of uniquely functionalized metabolites that can serve as pathway precursors (Figure 1.1) – namely those between 3-dehydroshikimate (3-DHS) and the terminal aromatic amino acid (AAA) products (i.e., tyrosine (Tyr), phenylalanine (Phe), and tryptophan (Trp)). Lying at a highly regulated metabolic node, chorismate serves as a particularly useful precursor to numerous aromatic pathways. Meanwhile, with conserved 4-hydroxyl functionality, phenolic products have typically been derived from the Tyr branch, including phenol (Kim et al. 2014), *p*-hydroxybenzoate (Williams et al. 2015), coumaric acid (Vargas-Tah et al. 2015) caffeic acid (Kang et al. 2012; Zhang and Stephanopoulos 2013), ferulic acid (Kang et al. 2012), *p*-hydroxystyrene (Qi et al. 2007), tyrosol (Satoh et al.

2012a), and hydroxytyrosol (Satoh et al. 2012b). Stemming from the Phe branch, pathways have been engineered to 2-phenylethanol (Kang et al. 2014; Koma et al. 2012), cinnamic acid (Vargas-Tah et al. 2015), styrene (McKenna and Nielsen 2011; McKenna et al. 2014), (*S*)-styrene oxide (McKenna et al. 2013), (*R*)-1,2-phenylethanediol (McKenna et al. 2013), (*S*)- and (*R*)-mandelic acid (Sun et al. 2011), and benzaldehyde (Kunjapur et al. 2014). Although comparatively fewer routes have to date been constructed from the Trp branch, notable examples include indigo dye (Berry et al. 2002) and catechol (though subsequently converted to muconic acid) (Sun et al. 2013).

1.3. Constructing Aromatic Biosynthesis Pathways from Robust Enzyme 'Parts'

With a vast and comprehensive 'toolkit' of known and/or putative enzyme chemistries – including those with high chemo-, regio-, and/or stereo-specificity – efficient construction of aromatic biosynthetic pathways is further supported by the availability of an abundance of useful enzyme 'parts'. In this case, inherent metabolic diversity draws from both anabolic and catabolic processes, co-evolved predominantly as a result of plant-microbe interactions. Within the rhizosphere and throughout the soil environment, for example, secondary aromatic metabolites synthesized by plant pathways (e.g., phenolic signaling and defense molecules, as well as lignin) (Douglas 1996; Herrmann 1995) are ultimately assimilated and catabolized via a series of microbial 'funneling pathways' (Linger et al. 2014). Adding to this are additional enzymes and pathways disparately evolved to detoxify (Beekrum et al. 2003; Stratford et al. 2007) and/or degrade (Diaz 2004) other, often xenobiotic, aromatic compounds.

In many cases, it is often hydroxyl group functionality at the 4- (or *p*-) position that is found to be a strong determinant of enzyme substrate specificity. For example,

while various Gram-positive bacteria have emerged as troves of broad specificity aromatic acid decarboxylases, said activities are typically limited only to substrates with 4-hydroxyl functionality (e.g., ferulic acid, caffeic acid, and others) (Tran et al. 2008). In contrast, meanwhile, phenylalanine ammonia lyases (PALs) from dicotyledonous plants (including those from *A. thaliana*), maintain a strict specificity for Phe and display little to no tyrosine ammonia lyase (TAL) activity (i.e., are incompatible with 4-hydroxyl functionalized substrates) (Cochrane et al. 2004). Interestingly, in one study it was shown that PAL vs. TAL activity could ultimately be controlled by changing single residue in the substrate binding pocket: a hydrophobic residue (Phe) gives rise to PAL activity, while a residue that promotes hydrogen bonding (His) results in TAL activity (Watts et al. 2006).

Such examples of protein engineering to successfully control substrate specificity are rare, however, and, as a result of such behaviors, a priori selection of

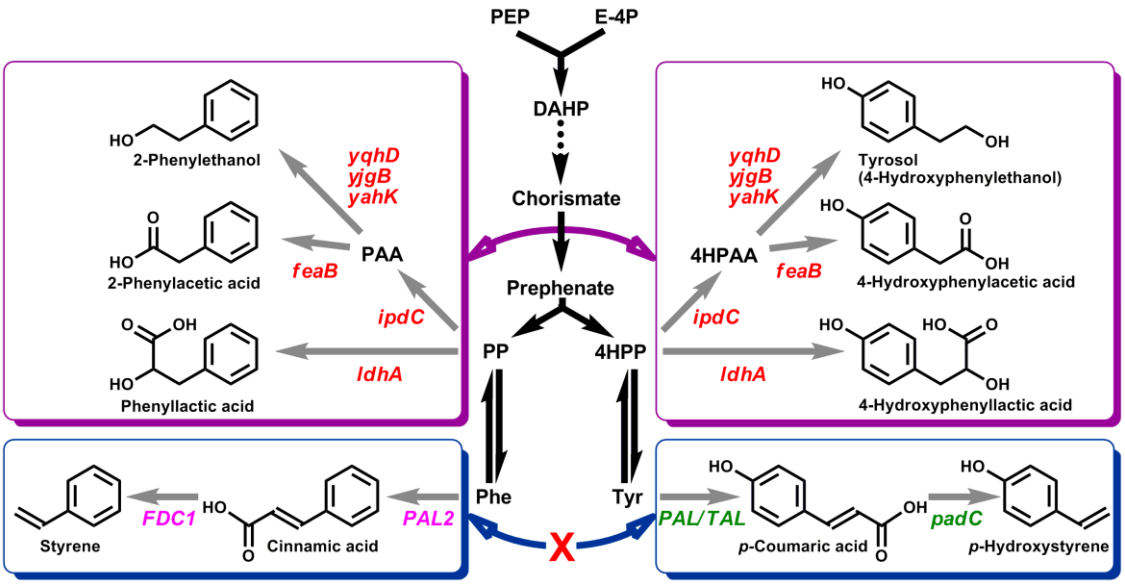


Figure1.2 Chorismate Derived Aromatic Biosynthesis.Enzyme 'parts' with broad substrate specificity have, in some cases, been exploited to allow engineered pathways to be leveraged for targeting additional products of interest. In other cases, however, biosynthesis of analogous products requires the construction of unique enzyme pathways.

candidate enzymes remains a challenge with respect to the engineering of aromatic pathways. An example of this is seen when comparing recent efforts to engineer pathways towards *p*-hydroxystyrene and styrene – two aromatics used extensively in plastics production (McKenna and Nielsen 2011; Qi et al. 2007). Despite involving analogous structures and requisite enzyme chemistries, their individual biosynthesis was only possible by engineering two unique pathways (*PAL/TAL* and *padC* vs. *PAL2* and *FDC1*, respectively), neither of which was found to display any significant mutual compatibility (Figure 1.2).

1.4. Exploiting Modularity to Enhance and Expand Aromatic Chemical Production

One effective and emerging approach in pathway engineering that has proven particularly useful for producing aromatic chemicals involves discretizing pathways into a series of complementary modules. Juminaga *et al.*, for example, applied this strategy to enhance the native production of tyrosine by optimizing relevant expression parameters (Juminaga et al. 2012). The tyrosine pathway was split at shikimate into upper and lower modules to enable efficient flux debottlenecking as a result of optimized relative plasmid copy number and promoter strength. The same approach has also been successfully applied to improve *E. coli*'s ability to synthesize (2S)-pinocembrin – a flavonoid of pharmaceutical interest derived from Phe (via cinnamic acid; Figure 1.1) – by as much as 56-fold (Wu et al. 2013).

In addition to promoting enhanced flux, modular approaches to pathway engineering further facilitate the ability by which additional aromatic products can be systematically targeted. For example, facile, step-wise extension of existing pathways can be performed to introduce additional product functionality. The authors have explored this prospect, for example, by demonstrating how the chiral aromatic fine

chemicals (*S*)-styrene oxide and (*R*)-1,2-phenylethanediol can both be synthesized by single-step extensions of the previously-engineered styrene pathway (McKenna et al. 2013). More significantly, by leveraging common intermediates, the 'plug and play' assembly of different yet complementary pathway modules allows metabolite flux to be easily redirected towards range of alternative end products. Wang *et al.*, for example, recently developed a library of mutually compatible 'biosynthetic bricks' for constructing modular pathways towards several phenylpropanoid derivatives (Wang et al. 2015). By this approach, a total of four phenylpropanoic acids (cinnamic acid, p-coumaric acid, caffeic acid, and ferulic acid), three stilbenoids (resveratrol, piceatannol and pinosylvin), and three curcuminoids (curcumin, bisdemethoxycurcumin and dicinnamoylmethane) were synthesized by *E. coli*.

Another excellent application of modular pathway engineering to enhance the bioproduction of aromatic chemicals is also seen in the recent works of Koma *et al.* (Koma et al. 2012). In said study, a suite of interrelated, heterologous pathways was engineered in *E. coli* to enable the individual biosynthesis of several unique aromatic alcohol and acid products (Figure 1.2). By employing the heterologous Ehrlich pathway, the flavor/fragrance compound 2-phenylethanol was first synthesized from endogenous phenylpyruvate. Then, following elimination of endogenous aldehyde reductase activity and over-expression of phenylacetaldehyde dehydrogenase, the common intermediate 2-phenylacetaldehyde was further leveraged to enable the alternative biosynthesis of 2-phenylacetate, a drug precursor. Meanwhile, phenylpyruvate was also alternatively and directly reduced to 2-phenyllactate using a lactate dehydrogenase with broad substrate specificity from *Cuprvidus necator*. Furthermore, in contrast to the above examples of rigid rules for natural substrate specificity surrounding 4-hydroxyl functionality, the same broad specificity pathway enzymes, were analogously applied to construct mirroring pathways that stemmed

instead from 4-hydroxyphenylpyruvate (4HPP; Figure 1.2). As a result, it was demonstrated that the same engineered pathways could be further leveraged to achieve the additional biosynthesis of phenolic analogs of the same three products.

1.5. Multiple Pathways, One Product

Another recent and notable development associated with the bioproduction of aromatic chemicals involves the engineering of multiple distinct routes to the same terminal product. Utilizing different precursors and/or associated chemistries, alternative pathways can provide effective options for circumventing inherent limitations of established routes. Recent examples of this trend have been seen in the case of both caffeic acid (Lin and Yan 2012) and 2-phenylethanol (Achmon et al. 2014; Masuo et al. 2015). However, just as in natural 'funneling pathways', the most significant examples of this trend to date have predominantly involved the engineering of pathways to phenolic products. For example, as originally demonstrated in *Pseudomonas putida* and more recently applied to *E. coli*, the biosynthesis of phenol – a versatile commodity chemical used widely in plastics production – has traditionally been achieved from endogenous Tyr via the action of tyrosine phenol lyase (TPL) (Kim et al. 2014). However, TPL is both reversible (actually favoring Tyr formation) and subject to feedback inhibition by phenol, and has accordingly been suggested as the principal limiting factor in phenol bioproduction (Wierckx et al. 2008). Recent efforts by Miao *et al.* have shown that TPL can be bypassed *en route* to phenol biosynthesis via an alternative pathway that stems instead from chorismate, proceeding via the intermediate *p*-hydroxybenzoate (Figure 1.3) (Miao et al. 2015a).

Catechol provides another example of a phenolic compound that has been targeted by multiple unique routes thanks to pathway engineering. Though typically explored as a precursor to muconic acid, catechol is itself useful for the synthesis of

fine chemicals, flavors/fragrances, and pharmaceuticals. While the classical pathway to catechol utilizes DHS as its immediate endogenous precursor, several unique and alternative routes have recently been reported (Averesch and Krömer 2014; Lin et al. 2014; Pugh et al. 2014; Sun et al. 2013; Sun et al. 2014), all of which stem instead from chorismate and thereby mitigate the auxotrophy limitations associated with original, DHS-derived pathway.

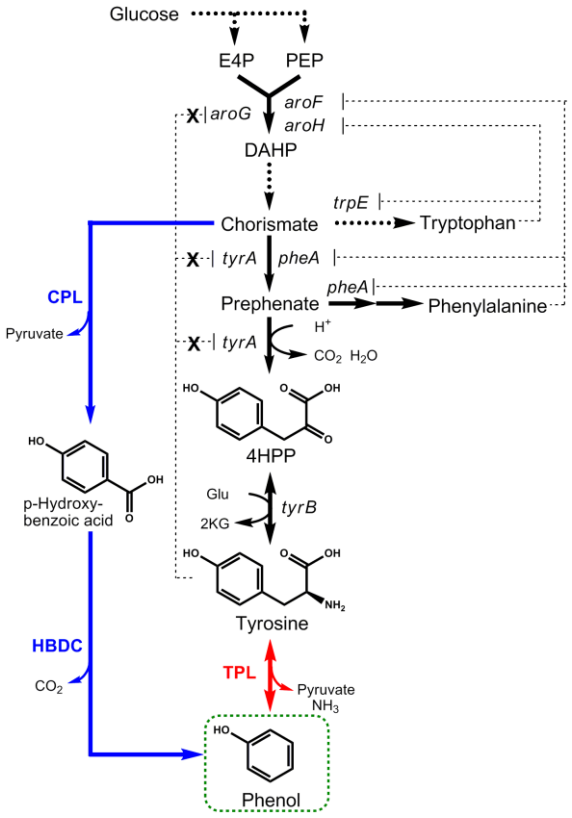


Figure 1.3 Phenol Biosynthesis Pathways.The biosynthesis of phenol has traditionally been achieved with the use of tyrosine phenol lyase (TPL). To circumvent inherent shortcomings of TPL, however, an alternative has been constructed comprised of chorismate pyruvate lyase (CPL) and 4-hydroxybenzoate decarboxylase (HBDC).

1.6. Inherent Challenges, Future Outlooks, and Conclusions

Whereas continued pathway engineering efforts will undoubtedly support a further expanding portfolio of renewable aromatic products, the ability to achieve economical production metrics for both existing and future targets will likely remain a challenge due to several inherent limitations. Toxicity is a common challenge for most aromatics, and one that limits achievable product titers (McKenna et al. 2013; Ramos et al. 2002). It may be possible o address this issue, however, by taking cues from nature. For example, many wild-type microbes have evolved inherent tolerances to a variety of aromatics (e.g., *Pseudomonas* sp.), and do so by fortifying their membrane structure and/or actively expelling species from within with the aid of efflux pumps (Isken and de Bont 1998; Ramos et al. 2002).

Meanwhile, among all proteinogenic amino acids, Tyr, Phe, and Trp biosynthesis carry the greatest metabolic costs, and, as a result, are consequently among the least abundant amino acids in most wild-type systems (Akashi and Gojobori 2002). Accordingly, significant host engineering is typically necessary to ensure an adequate supply of the requisite pathway precursor. While this problem has been addressed in several common laboratory strains, the same cannot be said with respect to other emerging and/or industrial hosts. Furthermore, aerobic conditions are typically required to provide adequate ATP needed for AAA biosynthesis, which, in addition to increasing production costs, also supports robust biomass production – a significant competitor for available nutrients that challenges the ability to achieve near-theoretical product yields. However, with only modest maximum theoretical glucose yields (e.g., 0.57, 0.57, and 0.44 mol/mol for Tyr, Phe, and Trp, respectively) (Kaleta et al. 2013), even meeting such a mark does not ensure economic viability of most prospective aromatic products. Accordingly, fine and specialty aromatic products are likely to

emerge as the first industrial successes while commodities, at least for now, are likely to face challenges with respect to commercialization.

Aromatic compounds represent a useful and versatile class of chemicals that have emerged as promising targets for metabolic pathway engineering. Although inherent limitations may challenge the microbial biosynthesis of all but fine chemical targets to economically-viable levels in the near term, several unique and enabling factors render these aromatic compounds as fertile ground for continued fundamental pursuits in pathway engineering.

1.7. Dissertation Organization

This dissertation is organized into five chapters. Chapter 1 provides a brief introduction to and motivation for the microbial biosynthesis of aromatic compounds, highlighting notable achievements and current limitations associated with engineering aromatic biosynthesis routes. Chapter 2 explores the engineering of two novel routes for the biosynthesis of phenol from glucose and offers a comparison to the established route. Further building upon these efforts, Chapter 3 investigates the development of four novel muconic acid (MA) biosynthesis pathways and provides similar comparisons with the established route. To maximize precursor assimilation in support of the biosynthesis of phenol, MA, and other potential bioproducts of interest, Chapter 4 introduces and explores the merits of a new metabolic engineering concept: synthetic 'metabolic funneling'. Finally, Chapter 5 offers an outlook on potential future work focused on phenotype control switch mechanisms for dynamic carbon flux regulation to further increase production metrics of phenol, MA, and other compounds. Taken together, this dissertation demonstrates renewable phenol and MA production exceeding previously reported efforts while identifying key opportunities for further increases to relevant production metrics.

2. ENGINEERING AND COMPARISON OF NON-NATURAL PATHWAYS FOR MICROBIAL PHENOL PRODUCTION

Abstract

The non-renewable petrochemical phenol is used as a precursor to produce numerous fine and commodity chemicals, including various pharmaceuticals and phenolic resins. Microbial phenol biosynthesis has previously been established, stemming from endogenous tyrosine via tyrosine phenol lyase (TPL). TPL, however, suffers from feedback inhibition and equilibrium limitations, both of which contribute to reduced flux through the overall pathway. To address these limitations, two novel and non-natural phenol biosynthesis pathways, both stemming instead from chorismate, were constructed and comparatively evaluated. The first proceeds to phenol in one heterologous step via the intermediate *p*-hydroxybenzoic acid, while the second involves two heterologous steps and the associated intermediates isochorismate and salicylate. Maximum phenol titers achieved via these two alternative pathways reached as high as 377 ± 14 and 259 ± 31 mg/L in batch shake flask cultures, respectively. In contrast, under analogous conditions, phenol production via the established TPL-dependent route reached 377 ± 23 mg/L, which approaches the maximum achievable output reported to date under batch conditions. Additional strain development and optimization of relevant culture conditions with respect to each individual pathway is ultimately expected to result in further improved phenol production.

This chapter was published as:

Thompson, B., Machas, M., Nielsen, D.R. Engineering and Comparison of Non-natural Pathways for Microbial Phenol Production (2015). *Biotech Bioeng,* 113(8):1745-1754.

2.1. Introduction

Phenol is an aromatic commodity chemical used widely as a versatile building block molecule for the synthesis of various plastics, phenolic resins, synthetic fibers, and most notably, bisphenol A (Weber et al. 2000). However, as all phenol is conventionally derived from various petrochemical feedstocks (including benzene, toluene, and propylene) (Weber et al. 2000), current production routes are non-renewable and thus unsustainable. To address this limitation, several researchers have explored alternative routes for phenol production directly from renewable resources, most notably biomass-derived sugars, through the engineering and use of microbial biocatalysts.

The first engineered pathway for phenol biosynthesis from glucose was reported by Wierckx *et al.* using *Pseudomonas putida* S12, a naturally solvent tolerant microbe, as the host platform (Wierckx et al. 2005). Phenol was synthesized from endogenous tyrosine by expression of a heterologous tyrosine phenol lyase (TPL), encoded by the gene *tpl* from *Pantoea agglomerans* (Figure 2.1). Random mutagenesis and antimetabolite selection were applied to deregulate native tyrosine biosynthesis, thus improving its endogenous availability and achievable phenol production metrics. Ultimately, phenol production from glucose was demonstrated in batch shake flask cultures with titers reaching 141 mg/L at a glucose yield of 0.034 g/g. Subsequent adoption of a fed-batch operating protocol then ultimately enabled phenol production at up to 471 mg/L. More recently, the same pathway was also reconstructed in *Escherichia coli*, in this case using *Pasteurella multocida* 36950 as the source of the *tpl* gene (Kim et al. 2014). A total of 18 different *E. coli* strains were evaluated as potential hosts, including each of K-12, B, and W strains as well as a K-12/B hybrid. In this case, tyrosine over-production was systematically engineered using synthetic small regulatory RNA (sRNA) to knock-down expression of two key

tyrosine biosynthesis pathway regulators (*tyrR* and *csrA*) while also over-expressing several other genes in the tyrosine biosynthesis pathway (namely, wild-type *tktA*, *ppsA*, *aroF*, *aroK*, and *tyrC*, along with feedback resistant mutants of *aroG* and *tyrA*) to overcome additional flux bottlenecks. Using *E. coli* BL21 as the host, phenol titers in shake flask cultures ultimately reached up to 419 mg/L under batch conditions, with fed-batch operation enabling production at up to 1.69 g/L at a glucose yield of 0.0026 g/g. For comparison, the maximum theoretical yield of phenol from glucose via said pathway has been reported as 0.38 g/g (Kim et al. 2014), as calculated using the *E. coli* genome-scale model EcoMBEL979 (Reed et al. 2003).

Despite these important results, the established phenol biosynthesis pathway is known to suffer from several fundamental shortcomings that limit its current efficacy and future prospects. First, as indicated by its positive change in Gibbs free energy due to reaction ($\Delta_r G'^{\circ}$ = +27.9 kJ/mol) in Figure 2.1, the reaction catalyzed by TPL is reversible and actually favors tyrosine formation. In fact, owing to this feature, TPL has long been exploited as an alternative enzymatic route for synthesizing tyrosine from exogenous phenol (Nagasawa et al. 1981; Para et al. 1985), as well as various non-canonical amino acids from other phenolic substrates (e.g., catechol) (Foor et al. 1993). Second, TPL is known to be subject to feedback inhibition which is caused by phenol at relatively low concentrations. For example, in the presence of just 94 mg/L phenol, it has been shown that TPL retains only 23% of its maximal activity (i.e., in the absence of phenol) (Wierckx et al. 2008). For these reasons, Wierckx *et al.* specifically pointed to TPL as the limiting factor for phenol production by *P. putida* (Wierckx et al. 2008). In the case of *E. coli*, these limitations are clearly manifested in the form of low pathway efficiency due to a metabolite flux bottleneck at the final step. More specifically, although Kim *et al.* were able to achieve phenol production at up to 419 mg/L in batch cultures, between 60 and 350 mg/L of unconverted tyrosine

was also observed to terminally co-accumulate in the culture medium (Kim et al. 2014). Though not specifically emphasized in their study, these outcomes were likely the result of low TPL activity.

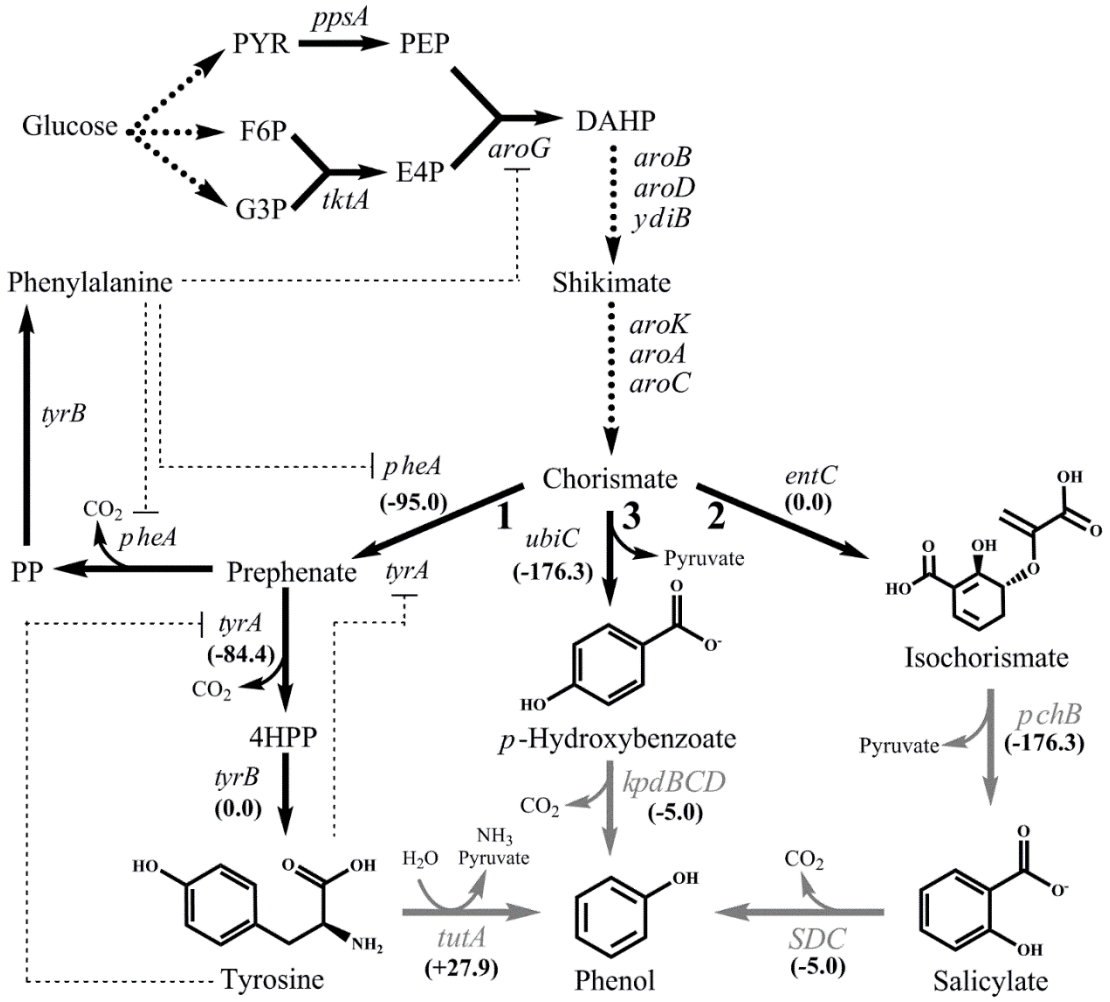


Figure 2.1 Engineered Phenol Biosynthesis.Three non-natural pathways engineered for phenol production from glucose. Black arrows represent enzyme steps native to *E. coli* whereas gray arrows are heterologous; bold dotted arrows represent multiple enzymatic steps. $\Delta_r G'^{\circ}$ is the change in Gibbs free energy due to reaction as determined using the online tool eQuilibrator (http://equilibrator.weizmann.ac.il) at a reference state of 25°C, pH 7, and ionic strength of 0.1 M. G3P: D-glyceraldehyde 3-phosphate. F6P: β-D-fructofuranose 6-phosphate. E4P: D-erythrose 4-phospahte. PYR: pyruvate. PEP: phosphoenolpyruvate. DAHP: 3-deoxy-D-arabino-heptulosonate-7-phosphate. PP: phenylpyruvate. 4HPP: 4-hydroxyphenylpyruvate.

In light of these limitations, it was postulated that the development of a TPL-independent biosynthesis pathway would constitute an effective approach towards improving the microbial production of renewable phenol. Accordingly, as illustrated in Figure 2.1, two novel and non-natural phenol biosynthesis pathways were proposed and investigated to this end. In contrast to the established route (Pathway 1), both novel pathways stem instead from chorismate as their immediate endogenous precursor. In the proposed Pathway 2, chorismate is first converted to isochorismate by isochorismate synthase, isochorismate is then converted to salicylate by isochorismate pyruvate lyase, and finally salicylate is transformed to phenol by salicylate decarboxylase. In the proposed Pathway 3, chorismate is first converted to *p*-hydroxybenzoate by chorismate pyruvate lyase before *p*-hydroxybenzoate is subsequently transformed to phenol by *p*-hydroxybenzoate decarboxylase. Similar to the established phenol pathway (Pathway 1), Pathway 3 involves the expression of a single heterologous enzyme, whereas Pathway 2 involves two heterologous steps (note: *E. coli* possesses a native chorismate pyruvate lyase, *ubiC*, and two isochorismate synthase genes, *entC* and *menF*). In addition, both Pathway 2 and 3 terminate via decarboxylation reactions which, unlike TPL, are rendered as essentially irreversible due to the release of CO_2 gas from the culture medium (Shen et al. 2011). Thus, in contrast to Pathway 1, greater conversion efficiency through said novel routes was expected as a result. With these factors in mind, the focus of the present study was to systematically construct and then comparatively evaluate the performance of the two novel phenol biosynthesis pathways with the goal of understanding if and how well their perceived inherent advantages can ultimately be translated into enhanced phenol biosynthesis relative to the established route.

2.2. Materials and Methods

2.2.1. Bacterial Strains and Media

All strains used in this study are listed in Table 2.1. Strains were routinely cultured in Luria-Bertani (LB) broth at 32ºC supplemented with 100 mg/L ampicillin, 35 mg/L kanamycin, and/or 34 mg/L chloramphenicol, as appropriate. To test for phenol production, relevant strains were cultured at 32ºC in a phosphate-limited minimal media composed of (concentrations in parentheses, all in g/L): $MgSO_4 \cdot 7H_2O$ (0.5), $(NH_4)_2SO_4$ (4.0), MOPS (24.7), KH_2PO_4 (0.3), K_2HPO_4 (0.7), $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ ($3.7 \cdot 10^{-4}$), H_3BO_3 ($2.5 \cdot 10^{-3}$), $CoCl_2 \cdot 6H_2O$ ($7.14 \cdot 10^{-4}$), $CuSO_4$ ($1.6 \cdot 10^{-4}$), $MnCl_2 \cdot 4H_2O$ ($1.6 \cdot 10^{-3}$), $ZnSO_4 \cdot 7H_2O$ ($2.88 \cdot 10^{-4}$), $FeCl_3$ ($5.0 \cdot 10^{-5}$), and glucose (20), supplemented with appropriate antibiotics. *E. coli* NEB10-beta (New England Biolabs (NEB), Ipswich, MA) was used for all cloning work and plasmid propagation. *Pseudomonas aeruginosa* PAO1 was obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany) and served as the genetic source for *pchB,* which encodes isochorismate synthase activity. *Citrobacter braakii* (ATCC 29063) was obtained from the American Type Culture Collection (ATCC, Manassas, VA) and served as the genetic source of *tutA*, which encodes TPL activity. *E. coli* BW25113 was obtained from the Coli Genetic Stock Center (CGSC, New Haven, CT) and served as the genetic source of *ubiC*, which encodes chorismate pyruvate lyase activity, in addition to *entC* and *menF*, which both individually encode isochorismate synthase activity. *Klebsiella pneumoniae* PZH572 (ATCC 25955) was obtained from the ATCC and served as the genetic source of *kpdBCD*, which encodes pHBA decarboxylase activity. *Bacillus subtilis* (ATCC 31002) was obtained from the ATCC and served as the genetic source of *bsdBCD*, which also encodes *p*-hydroxybenzoate decarboxylase activity.

2.2.2. Plasmid Construction

All plasmids constructed and/or used in this study are listed in Table 2.1. Candidate genes were PCR amplified using a BioRad iCycler system with Q5 High-Fidelity DNA Polymerase (NEB) using standard protocols and custom DNA oligonucleotide primers (see Appendix A, Table A1) synthesized by Integrated DNA Technologies (Coralville, IA). Genomic DNA (gDNA) was prepared from cell cultures using the ZR Fungal/Bacterial DNA MiniPrep (Zymo Research, Irvine, CA) according to manufacturer protocols. Amplified linear DNA was purified using the Zymo Research DNA Clean & Concentrator kit (Zymo Research). Purified DNA fragments and plasmids were digested with select restriction endonucleases (NEB) according to manufacturer protocols. Digested DNA fragments were gel purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research) and ligated with T4 DNA Ligase (NEB) according to manufacturer protocols. Ligation reactions were transformed into chemically competent *E. coli* NEB10-beta and selected for by plating on LB solid agar containing appropriate antibiotics. Transformant pools were screened with the use colony PCR and restriction digest mapping.

To construct the plasmids pTutA, pKpd, and pBsd, *tutA*, *kpdBCD*, and *bsdBCD* were PCR amplified from *C. braakii*, *K. pneumoniae*, and *B. subtilis* gDNA, respectively, and individually inserted into pTrcCOLAK. To construct plasmid pSDC, *SDC* was PCR amplified from a codon-optimized and synthesized (GenScript, Piscataway, NJ) variant of *SDC* from *Trichosporon moniliiforme* and inserted into pTrcCOLAK. To construct plasmid pUbiC-Kpd, *ubiC* was PCR amplified from *E. coli* BW25113 gDNA and inserted to pKpd. To construct plasmid pSDC-PchB-EntC, *pchB* and *entC* were PCR amplified from *P. aeruginosa* PAO1 and *E. coli* BW25113 gDNA, respectively, and inserted to pSDC. Plasmids pS3 and pY3 were gifts from Jay Keasling (Addgene plasmid #50596 and #50606, respectively). To construct pY3c, the P_{trc}-*aroC*-*aroA*-*aroL* and *lacI*-Amp^r-

p15A cassettes were both individually PCR amplified using pY3 as template and subsequently combined by circular polymerase extension cloning (CPEC) (Quan and Tian 2011).

2.2.3. Strain Construction

E. coli BW25113 was individually transformed with pTutA, pSDC-PchB-EntC, and pUbiC-Kpd, resulting in the construction of strains TYR01, SAL01, and HBA01, respectively. Strains TYR01, SAL01, and HBA01 were each individually transformed with pS3, resulting in the construction of strains TYR02, SAL02, and HBA02, respectively. Strains TYR02, SAL02, and HBA02 were each individually transformed with pY3, resulting in the construction of strains TYR03, SAL03, and HBA03, respectively. Strains SAL02 and HBA02 were each individually transformed with pY3c, resulting in the construction of strains SAL04 and HBA04, respectively. *E. coli* NST74 was individually transformed with pSDC-PchB-EntC and pUbiC-Kpd, resulting in the construction of strains SAL05 and HBA05, respectively. *E. coli* NST74 Δ*pheA* was individually transformed with pSDC-PchB-EntC and pUbiC-Kpd, resulting in the construction of strains SAL06 and HBA06, respectively. *E. coli* NST74 Δ*pheA* was co-transformed with pY3 and pTutA, resulting in the construction of TYR06.

Table 2.1 Strains and Plasmids for Phenol Production

Strains	Description	Source
E. coli NEB 10-beta	*Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)*	NEB
C. braakii ATCC 29063	Source of *tutA*	ATCC
P. aeruginosa PAO1	Source of *pchB*	DSMZ
E. coli BW25113	Source of *ubiC* and *entC*	CGSC
K. pneumoniae PZH572	Source of *kpdBCD*	ATCC
B. subtilis ATCC 31002	Source of *bsdBCD*	ATCC
E. coli NST74	*aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*	ATCC
E. coli NST74 Δ*pheA*^{fbr}	Chromosomal deletion of *pheA*^{fbr} in *E. coli* NST74	Pugh et al. (2014)
SAL01	*E. coli* BW25113 pSDC-PchB-EntC	This study
SAL02	SAL01 pS3	This study
SAL03	SAL02 pY3	This study
SAL04	SAL03 pY3c	This study
SAL05	*E. coli* NST74 pSDC-PchB-EntC	This study
SAL06	*E. coli* NST74 Δ*pheA* pSDC-PchB-EntC	This study
HBA01	*E. coli* BW25113 pUbiC-Kpd	This study
HBA02	HBA01 pS3	This study
HBA03	HBA02 pY3	This study
HBA04	HBA02 pY3c	This study
HBA05	*E. coli* NST74 pUbiC-Kpd	This study
HBA06	*E. coli* NST74 Δ*pheA* pUbiC-Kpd	This study
TYR01	*E. coli* BW25113 pTutA	This study
TYR02	TYR01 pS3	This study
TYR03	TYR02 pY3	This study
TYR06	*E. coli* NST74 Δ*pheA* pTutA, pY3	This study

Plasmid	Description	Source
pTrcCOLAK	ColA ori; Kan^r, *lacIq*, P_{trc}	McKenna et al. (2013)
pTutA	*tutA* from *C. braakii* ATCC 29063 inserted into pTrcCOLAK	This study
pSDC	codon-optimized variant of SDC from *T. moniliiforme* inserted into pTrcCOLAK	This study
pKpd	*kpdBCD* from *K. pneumoniae* PZH572 inserted into pTrcCOLAK	This study
pBsd	*bsdBCD* from *B. subtilis* ATCC 31002 inserted into pTrcCOLAK	This study
pSDC-PchB-EntC	*entC* from *E. coli* BW25113 and *pchB* of *P. aeruginosa* PAO1 inserted into pSDC	This study
pUbiC-Kpd	*ubiC* from *E. coli* BW25113 inserted into pKpd	This study
pS3	pBBR1 ori; Cm^r, *lacI* $P_{lac-UV5}$-*aroE-aroD-aroB*^{op}-*aroG*^{fbr}-*ppsA-tktA*	Juminaga et al. (2012)
pY3	p15A ori; Amp^r, *lacI*, $P_{lac-UV5}$-*tyrB-tyrA*^{fbr}-*aroC* T1-P_{trc}-*aroA-aroL*	Juminaga et al. (2012)
pY3c	p15A ori; Amp^r, *lacI*, $P_{lac-UV5}$-*aroC* T1-P_{trc}-*aroA-aroL*	This study

2.2.4. Assaying Candidate Enzyme Activity Using Whole Resting Cells

The functionality of select candidate pathway enzymes was analyzed through a series of *in vivo* resting cell assays. *E. coli* BW25113 was individually transformed with each of pTutA, pKpd, pBsd, and pSDC. Overnight seed cultures were prepared and used to inoculate (1% inoculum) 50 mL of LB in 250 mL baffled shake flasks supplemented with 35 mg/L kanamycin. Shake flask cultures were induced by addition of 10 µL 1M IPTG upon reaching an optical density at 600 nm (OD_{600}) ~ 0.5. Cells were cultured overnight (~12 h) at 32ºC and 200 RPM, and centrifuged at 3,000 x *g* for 5 min the following morning. Cell pellets were rinsed once with phosphate buffered saline (PBS) solution before then being re-suspended in 50 mL PBS in a 250 mL baffled shake flasks at an OD_{600} of ~4.0. Tyrosine was added to an initial concentration of 2.5 mM, while *p*-hydroxybenzoate, and salicylate were added to initial concentrations of 2.5 and 5 mM before then incubating said resting cell cultures at 32ºC with shaking at 200 RPM for up to 48 h. Aqueous samples (0.5 mL) were periodically removed for analysis of substrate and product concentrations by HPLC, as described below.

2.2.5. Biosynthesis of Phenol from Glucose in Shake Flask Cultures

Overnight seed cultures of phenol producing strains were prepared and used to inoculate 50 mL MM1 minimal media (1% inoculum), supplemented with 20 g/L glucose and appropriate antibiotics; NST74 Δ*pheA* strains were supplemented with 100 mg/L phenylalanine to account for auxotrophic conditions. Cultures were grown in 250 mL baffled shake flasks at 32ºC to promote shikimic acid pathway flux (Tribe 1987) with shaking at 200 RPM to an OD_{600} ~ 0.7, at which point 10µL of 1M IPTG was added. Culturing then continued under the same conditions for an additional 72 hours, with periodic sampling performed for metabolite analysis by HPLC, as described below.

2.2.6. HPLC Analysis

Prior to HPLC analysis, samples were diluted with 1 N HCl and incubated at 55°C for 30 minutes, followed by centrifugation at 11,000 x *g* for 3 minutes and subsequent supernatant transfer to glass HPLC vials. Metabolite analysis was performed using a Hewlett Packard 1100 series HPLC system. Separation of phenylalanine, *p*-hydroxybenzoate, salicylate, and phenol was achieved using a reverse-phase Hypersil GOLD aQ C18 column (4.6 mm x 150 mm; Thermo Fisher, USA) operated at 45°C with a constant flow rate of 0.8 mL/min consisting of 85% 5 mM sulfuric acid and 15% acetonitrile. The eluent was monitored using a diode array detector (DAD) set at 215 nm for phenylalanine and salicylate, 260 nm for *p*-hydroxybenzoate, and 275 nm for phenol. Separation of tyrosine was achieved with the same column using a constant flow rate of 0.5 mL/min at 30°C with a gradient of water (A) methanol/0.1% TFA (B) as follows: 5% B from 0 to 8 min, 5% to 40% B from 8 to 13 min, 40% B from 13 to 16 min, 40 to 5% B from 16 to 21 min, 5% B for 10 min. The eluent was monitored using a DAD set at 215 nm. Glucose analysis was performed using the same HPLC system equipped with a RID detector and an Aminex HPX-87H column (BioRAD, Hercules, CA) operated at 35ºC. The column was eluted with a constant flow rate of 0.55 mL/min 5 mM H_2SO_4. External calibrations were used to quantify each metabolite of interest.

2.2.7. Elementary Flux Mode Analysis

Elementary flux modes (EFMs) were computed in MATLAB R2014b (MathWorks, Natick, MA) using the freely available program EFMTool 4.7.1 (Terzer and Stelling 2008), available at: http://www.csb.ethz.ch/tools/efmtool. An *E. coli* stoichiometric network was adapted from the recent works of Averesch and Kromer (Averesch and Krömer 2014). Developed based on input form literature (Kromer et al. 2006) and

various pathway databases (Caspi et al. 2008; Kanehisa and Goto 2000; Kanehisa et al. 2014), said network was originally used to investigate factors influencing the production of muconic acid, which is similarly derived from shikimic acid precursors. Here, however, reactions associated with muconic acid production were replaced with those necessary for phenol production, according to each of the three pathways of interest (Figure 2.1). The network (see Appendix A, Figure A3a-c) was comprised of reactions involved in central carbon metabolism, specifically including glycolysis, Entner-Doudoroff Pathway, pentose phosphate pathway, TCA cycle, glyoxylate cycle, glutamtate/glutamine interconversion, anaplerosis, and the electron transport chain. Biomass production is also accounted for, occurring from precursors derived from central metabolism. Amino acid biosynthesis, including aromatic amino acids, was not included (except in the case of tyrosine production as the precursor for Pathway 1).

2.3. Results and Discussion

2.3.1. Identifying and Characterizing the Requisite Pathway Enzymes

To circumvent the aforementioned inherent limitations associated with the previously established TPL-dependent phenol biosynthesis pathway (i.e., Pathway 1), two novel pathways (i.e., Pathways 2 and 3) were proposed, systematically constructed, and comparatively evaluated. Investigation of the chemistries involved in each case reveals that both Pathway 2 and 3 are predicted to benefit from a 16% greater net thermodynamic driving force (Figure 2.2) relative to Pathway 1. More specifically, from the last common precursor chorismate, the net change in Gibbs free energy due to reaction ($\Delta_r G'^{\circ}_{net}$) through both Pathways 2 and 3 is predicted to be -181.3 kJ/mol, compared to just -151.5 kJ/mol through Pathway 1 (determined using the online tool eQuilibrator (http://equilibrator.weizmann.ac.il) at a reference state of 25^{o}C, pH 7, and ionic strength of 0.1 M). All three pathways are predicted to support

nearly equivalent theoretical yields of phenol from glucose (in each case ~75%), as determined via elementary flux mode analysis (see Appendix A, Table A2 and Figures A1-A3). Overall, this outcome is not surprising since all three pathways are derived from chorismate as the common endogenous precursor and carbon is equally conserved in each case (with one CO_2 and one pyruvate removed downstream of chorismate in each pathway;2.1).

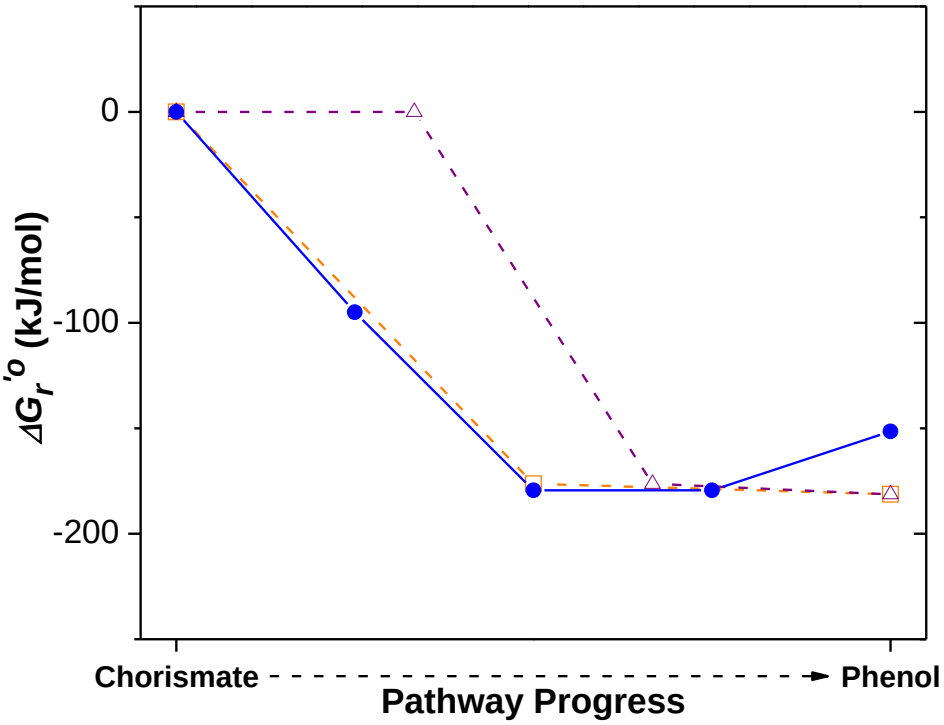


Figure 2.2 Gibbs Free Energy.Comparing the change in Gibbs free energy due to reaction with progress through each of the three pathways (Pathway 1, open circles-solid line; Pathway 2, open squares-dotted line; Pathway 3, open diamonds-dashed line) from chorismate (the last common precursor) to phenol. $\Delta_rG'°$ is the change in Gibbs free energy due to reaction as determined using the online tool eQuilibrator at a reference state of 25°C, pH 7, and ionic strength of 0.1 M (http://equilibrator.weizmann.ac.il).

To construct Pathway 2, it was first necessary to identify effective candidates displaying each of isochorismate synthase, isochorismate pyruvate lyase, and salicylate decarboxylase activities. *E. coli* possesses two native forms of isochorismate

synthase, encoded by *menF* and *entC*, which function as part of the menaquinone and enterobactin biosynthesis pathways, respectively (Dahm et al. 1998). Menaquinone is an essential electron carrier used during anaerobic growth (Guest 1977), and thus native *menF* expression has been reported to occur only under anaerobic conditions (Buss et al. 2001). Enterobactin is required for iron transport, and thus native *entC* expression primarily occurs in the absence of iron (Kwon et al. 1996). As neither environment represents the culture conditions of interest, it was determined that basal levels of isochorismate synthase activity would be insufficient for supporting phenol production in this study. As purified enzyme assays indicate that EntC possesses higher activity (K_m = 11.93 μM, k_{cat} = 2.12 s^{-1}) than MenF (K_m = 6.75 μM, k_{cat} = 0.13 s^{-1}) (Lin et al. 2013), *entC* was selected for over-expression as part of Pathway 2. Meanwhile, isochorismate pyruvate lyase activity has been identified as part of various secondary metabolite and siderophore biosynthesis pathways in a variety of bacterial species, and most notably in *Pseudomonas aeruginosa* where it is encoded by *pchB* (Gaille et al. 2002; Lamb 2011). As prior studies have already demonstrated that *pchB* expression in *E. coli* enables the efficient conversion of isochorismate to salicylate, with k_{cat} estimates as high as 11.2 s^{-1} (Lin et al. 2013), *pchB* was accordingly selected for over-expression as part of Pathway 2. Finally, salicylate decarboxylase activity has to date only been reported in the soil fungi *Trichosporon moniliiforme* WU-0401 (Kirimura et al. 2010). Encoded by *SDC*, salicylate decarboxylase functions as the first step in the native salicylate biodegradation pathway (Iwasaki et al. 2010). As its recombinant function in *E. coli* has not before been demonstrated, a codon-optimized variant of *SDC* was synthesized, cloned, and expressed in SAL01 resting cells in order to confirm and assess its *in vivo* function and activity. As seen in Figure 2.3, in addition to a maximum specific rate of 0.16 ± 0.02 mmol g^{-1} h^{-1}, approximately 20% and 11% of the initial 2.5 mM (345 mg/L) and 5 mM (690 mg/L) salicylate was

converted to phenol within 12 h, respectively. Based on these results, *SDC* was chosen for the terminal step in Pathway 2.

Construction of Pathway 3 first required identification of genes encoding chorismate pyruvate lyase and *p*-hydroxybenzoate decarboxylase activities. *E. coli* possesses a native copy of chorismate pyruvate lyase, encoded by *ubiC*, which functions as part of the ubiquinone biosynthesis pathway (Siebert et al. 1994). Although native biosynthesis of *p*-hydroxybenzoate biosynthesis is highly regulated in *E. coli* (Gallagher et al. 2001), the authors have previously shown that over-expression of native *ubiC* in a Δ*pheA* mutant of *E. coli* NST74 (previously engineered phenylalanine over-producing strain (Tribe 1987)) enabled the production of up to 277 mg/L *p*-hydroxybenzoate in shake flask cultures (Pugh et al. 2014). Based on this prior result, *ubiC* was selected for over-expression as the first step in Pathway 3. Phenol can be produced from *p*- hydroxybenzoate via *p*-hydroxybenzoate decarboxylase, an enzyme common to numerous soil microbes that functions most commonly as a key step in anaerobic aromatic biodegradation (Lupa et al. 2005). However, whereas numerous *p*-hydroxybenzoate decarboxylases identified to date are known to exhibit partial or even complete loss of functionality when expressed under aerobic environments (Lupa et al. 2005; Matsui et al. 2006), notable exceptions have been found in the case of oxygen-insensitive forms from *B. subtilis* and *K. pneumoniae* (encoded by *bsdBCD* and *kpdBCD*, respectively). Both *p*-hydroxybenzoate decarboxylase encoding candidate genes were cloned and screened with respect to their relative activity in whole resting cells of *E. coli* BW25113 pBsd and *E. coli* BW25113 pKpd. As seen in Figure 2.3B, when exogenous *p*-hydroxybenzoate was supplemented at 2.5 mM (345 mg/L) and 5 mM (690 mg/L), near stoichiometric conversion was achieved at both substrate levels within 12 h in the case of *E. coli* BW25113 pKpd, whereas overall conversion by *E. coli* BW25113 pBsd was significantly

less (~68%) in the same time (note: full stoichiometric conversion was ultimately achieved by 24 h in both cases; data not shown). Furthermore, maximum specific rates of *p*-hydroxybenzoate decarboxylation were also measured as 3.24 ± 0.4, 4.3 ± 0.07, and 8.9 ± 0.12 mmol g^{-1} h^{-1} for *E. coli* BW25113 pBsd, *E. coli* BW25113 pKpd with 2.5 mM pHBA fed, and *E. coli* BW25113 pKPD with 5 mM pHBA fed, respectively. Consequently, *kpdBCD* was selected for use in constructing Pathway 3.

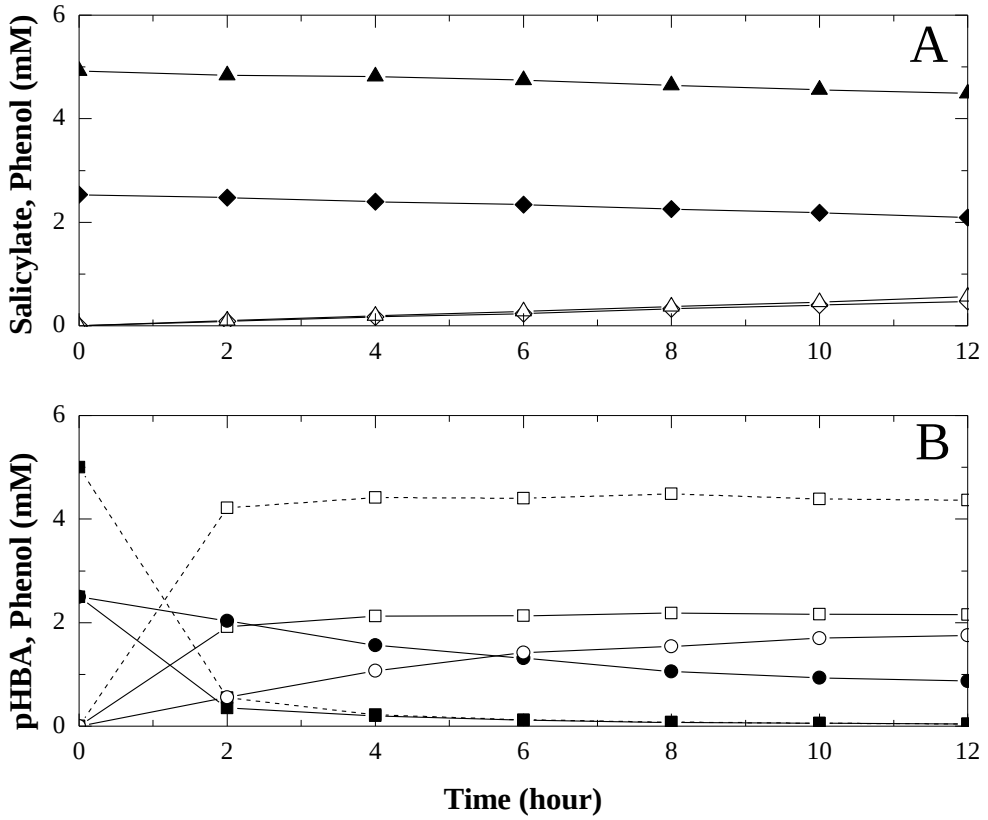


Figure 2.3 Recombinant Enzyme Activity.Screening candidate pathway enzymes in whole resting cell assays. (A) Conversion of 2.5 mM salicylate (solid diamond) to phenol (open diamond) and 5 mM salicylate (solid triangle) to phenol (open triangle) by *E. coli* BW25113 pSDC. (B) Conversion of 2.5 mM p-hydroxybenzoate (pHBA; solid circle) to phenol (open circle) by *E. coli* BW25113 pBsd, 2.5 mM *p*¬hydroxybenzoate (pHBA; solid square, solid line) to phenol (open square, solid line) and 5 mM pHBA (solid square, dashed line) to phenol (open square, dashed line) by *E. coli* BW25113 pKpd. Error bars represent one standard deviation from triplicate experiments.

2.3.2. Phenol Pathway Construction in Wild-type *E. coli*

Following identification of the suitable gene candidates for all of the required steps, full pathways were next assembled to test for phenol biosynthesis directly from glucose. To establish a performance baseline, wild-type *E. coli* BW25113 was used as the initial host background of interest. In addition to Pathways 2 and 3, Pathway 1 was also constructed to provide a fair comparison against the established route for phenol biosynthesis. In this case, *tutA* from *C. braakii* was selected to provide the needed TPL activity and its recombinant function and activity was first confirmed via a whole resting cell assay. In this case, 2.5 mM (453 mg/L) exogenous tyrosine was converted to phenol by *E. coli* BW25113 pTutA within 12 h at a maximum specific rate of 16.6 ± 0.206 mmol g^{-1} h^{-1} (see Appendix A, Figure A2). With all three pathways constructed, phenol biosynthesis directly from glucose was next evaluated. For Pathway 1, initial phenol titers by strain TYR01 reached 39 ± 10 mg/L in 72 h at a glucose yield of 2.0 ± 0.03 mg/g (Figure 2.4). Meanwhile, in the case of Pathway 2, higher initial phenol titers were achieved by strain SAL01, reaching 109 ± 16 mg/L in 72 h at a glucose yield of 5.4 ± 0.3 mg/g (Figure 2.4). Most notably, in the case of Pathway 3, strain HBA01 produced up to 259 ± 31 mg/L phenol in 72 h at a glucose yield of 12.9 ± 1.1 mg/g (Figure 2.4). It is important to note that no accumulation of any of the intermediate metabolites associated with either Pathway 2 or 3 were detected in the culture medium at any time (data not shown), indicating that no significant flux bottlenecks were experienced. In a wild-type host background, higher phenol titers achieved by Pathways 2 and 3 likely reflect the fact that chorismate (the immediate precursor for Pathways 2 and 3) is more prevalent than tyrosine (the immediate precursor for Pathway 1). Also derived from chorismate, tyrosine biosynthesis is additionally subjected to strong feedback regulation at the bifunctional chorismate mutase/prephenate dehydrogenase, encoded by *tyrA*, both by allosteric

inhibition (Lütke-Eversloh and Stephanopoulos 2005) as well as by dual transcriptional regulator TyrR (Pittard et al. 2005).

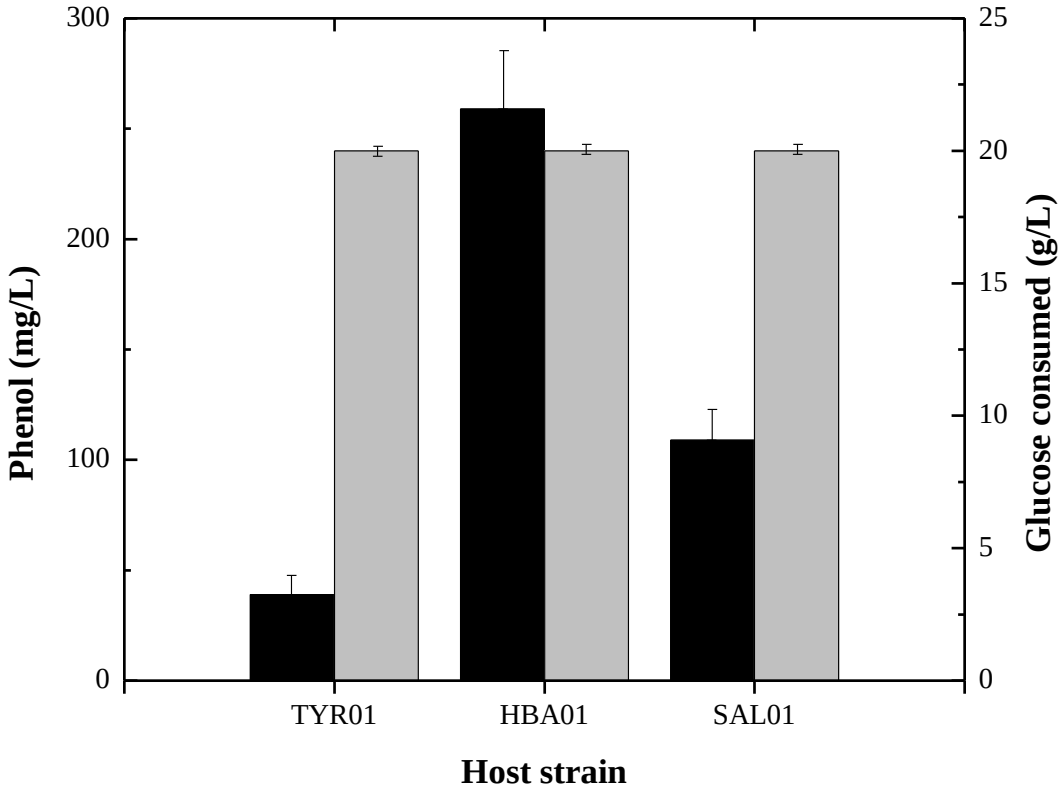


Figure 2.4 Wild-type Phenol Production.Maximum phenol produced (black) via Pathways 1, 2, and 3 using wild-type *E. coli* BW25113 as host background with glucose consumed (gray) after 72 hours of culturing in MM1 minimal media supplemented with 20 g/L glucose. Error bars represents one standard deviation from triplicate experiments.

2.3.3. Increasing Intracellular Availability of Precursor Chorismate in Support of Phenol Production

Although functionality of both Pathways 2 and 3 was successfully demonstrated, there remained significant room for improvement with respect to achievable phenol titers and yields. As no clear flux bottlenecks were observed through either pathway, it was hypothesized that this was due to the limited intracellular availability of chorismate in

the wild-type host. Subsequent strain engineering efforts were accordingly focused at improving metabolite flux towards and preserving the availability of this required precursor. The flux of metabolites entering the shikimic acid pathway is initially regulated by feedback inhibition of each of the three 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase isoenzymes encoded by *aroG*, *aroF*, and *aroH* (as controlled by phenylalanine, tyrosine, and tryptophan, respectively). Among these, AroG has been reported to account for ~80% of the total DAHP synthase activity in *E. coli* (Kikuchi et al. 1997). Although its wild-type activity is completely inhibited in the presence of only 16.5 mg/L (0.1 mM) phenylalanine (Rodriguez et al. 2014), AroG feedback resistant (*fbr*) mutants have been isolated with the ability to tolerate up to 1650 mg/L (10 mM) phenylalanine (Kikuchi et al. 1997). Simply over-expressing an $aroG^{fbr}$ mutant in *E. coli* has been shown to result in a 2.9-fold increase in DAHP production (Floras et al. 1996). Furthermore, in a study focused on enhancing shikimate biosynthesis by *E. coli*, Juminaga *et al.* also reported that, in addition to over-expressing $aroG^{fbr}$, further flux enhancements could be achieved by also over-expressing wild-type copies of additional genes in the upper shikimic acid pathway (specifically *aroE, aroD,* and *aroB*) along with select pentose phosphate pathway genes (*ppsA* and *tktA*) (Juminaga et al. 2012). In this case, said modifications were collectively implemented via the engineering and introduction of plasmid pS3 (Figure 2.5). A second plasmid module (pY3) was also constructed and used in said study to similarly overexpress the remaining genes in the lower shikimic acid pathway (i.e., shikimate through tyrosine; specifically *aroL*, *aroA*, *aroC*, *tyrB*, and a tyrosine feedback resistant mutant of *tyrA* ($tyrA^{fbr}$); Figure 2.5). Using these two plasmid modules, efficient flux debottlenecking and fine-tuning of relative expression levels of all shikimic acid pathway genes was achieved, ultimately enabling the production of over

2.2 g/L tyrosine from glucose at a yield of 0.44 g/g (80% the theoretical maximum yield) (Juminaga et al. 2012).

The effect of enhancing chorismate biosynthesis on phenol production was first investigated via introduction of pS3 into SAL01, HBA01, and TYR01. The resulting strains, SAL02, HBA02, and TYR02, respectively, were analogously evaluated in terms of phenol production, the results of which are compared in Figure 2.5. In the case of Pathway 2, phenol production by SAL02 ultimately reached 130 ± 37 mg/L at a glucose yield of 10.6 ± 2.6 mg/g, representing 19 and 96% increases over SAL01, respectively. Said strategy was similar effective for enhancing phenol production via Pathway 1, with maximum titers by TYR02 reaching 77 ± 9 mg/L at a glucose yield of 3.9 ± 0.01 mg/g, representing 97 and 95% increases over strain TYR01, respectively. However, in the case of Pathway 3, it was interestingly found that enhancing chorismate availability by this strategy had the opposite effect, with achievable phenol titers by HBA02 decreasing to just 193 ± 10 mg/L, despite a modest (~20%) increase in yield. At present, the specific reason underlying this observed flux decrease in the case of Pathway 3 remains unclear.

Although clearly providing initial flux enhancements towards chorismate (and ultimately phenol), pS3 includes only those upper shikimic acid pathway genes leading to shikimate, it was thus possible that the one of the remaining steps (i.e., *aroL*, *aroA*, or *aroC*) was still limiting en route to chorismate. To provide additional copies of key pathway enzymes downstream of shikimate and further increase flux towards chorismate (and ultimately tyrosine in the case of Pathway 3), the previously engineered plasmid pY3 was next introduced into each of TYR02, SAL02, and HBA02. The resultant strains (TYR03, SAL03, and HBA03) were again evaluated in terms of their phenol production abilities, the results of which are summarized in Figure 2.5. In the case of Pathway 1, phenol titers by strain TYR03 reached 304 ± 12 mg/L at a

glucose yield of 15.3 ± 0.1 mg/g, representing 679 and 665% increases over TYR01, respectively. However, as in prior studies employing the TPL-dependent pathway, it should be noted that residual tyrosine also remained at the completion culture, in this case reaching up to 142 ± 30 mg/L (Figure 2.5). Meanwhile, whereas ~100% increases in both titer and yield were observed in the case of Pathway 2 (between SAL03 and SAL01), a reduced titer was again observed for Pathway 3; in this case a 98% decrease between HBA03 and HBA01.

In addition to phenol, significant co-accumulation of tyrosine was also now observed in the case of Pathway 2 and 3 for strains carrying pY3 (with as much as 672 mg/L tyrosine detected at the completion of SAL03 cultures; Figure 2.5). However, in this case, tyrosine accumulation was not the result of an inefficient and/or reversible pathway enzyme, but rather resulted from the undesired co-expression of $tyrA^{fbr}$ on pY3. Thus, using pY3 as a template, an additional plasmid module (pY3c, Figure 2.5) was engineered for over-expression of *aroL*, *aroA*, and *aroC* alone by truncating the expression cassette to remove both *tyrB* and $tyrA^{fbr}$. However, as seen in Figure 2.5, anticipated gains in phenol production due to preserving chorismate availability in SAL04 and HBA04 (harboring both pS3 and pY3c) were not ultimately realized. In addition to a lack of improvement in phenol production, it was also noted that net cell growth by both SAL04 and HBA04 was significantly reduced (by as much as 65%) relative to SAL01 and HBA01, respectively (data not shown). It was hypothesized the increased metabolic burden (Birnbaum and Bailey 1991), caused by the need to simultaneously maintain and express three pathway-encoding plasmids (including two for precursor production alone) was a possible cause of observed reductions in cell fitness and limited overall phenol production. Accordingly, alternative host strains with the requisite, chromosomally-integrated mutations for enhancing precursor supply were lastly investigated as phenol production platforms.

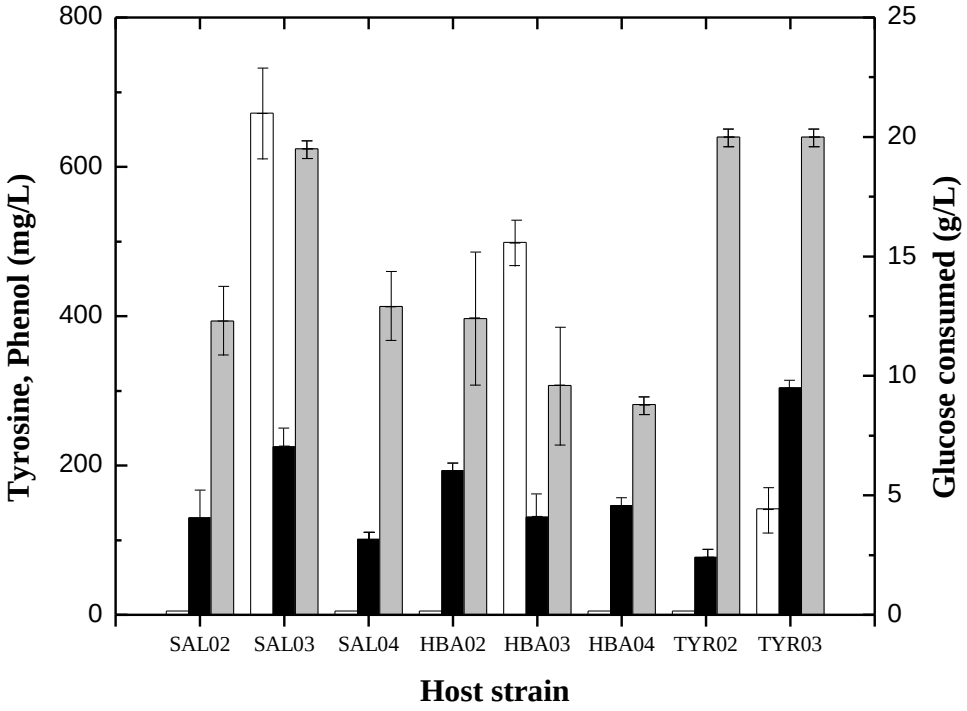


Figure 2.5 Host Strain Engineering to Support Phenol Production.Effects of different host strain engineering strategies designed to increase chorismate availability. Residual tyrosine (white), phenol produced (black), and glucose consumed (gray) following 72 hours of culturing in MM1 minimal media supplemented with 20 g/L glucose. Error bars represents one standard deviation from triplicate experiments.

2.3.4. Further Strain Engineering in Support of Phenol Biosynthesis

Previously engineered for phenylalanine over-production, *E. coli* NST74 (Tribe 1987) was next investigated as a potential host platform for phenol production by each of Pathways 1-3 (Figure 2.6). Although phenylalanine biosynthesis does not directly support phenol production by any of the three routes, NST74 carries the desired *aroG*^{fbr} mutant, along with several other key mutations (see Table 2.1) stably on its chromosome (Tribe 1987). While modest phenol titer and yield improvements were observed in the case of SAL05 versus SAL04 (see Tables 3 and 4), significant levels of

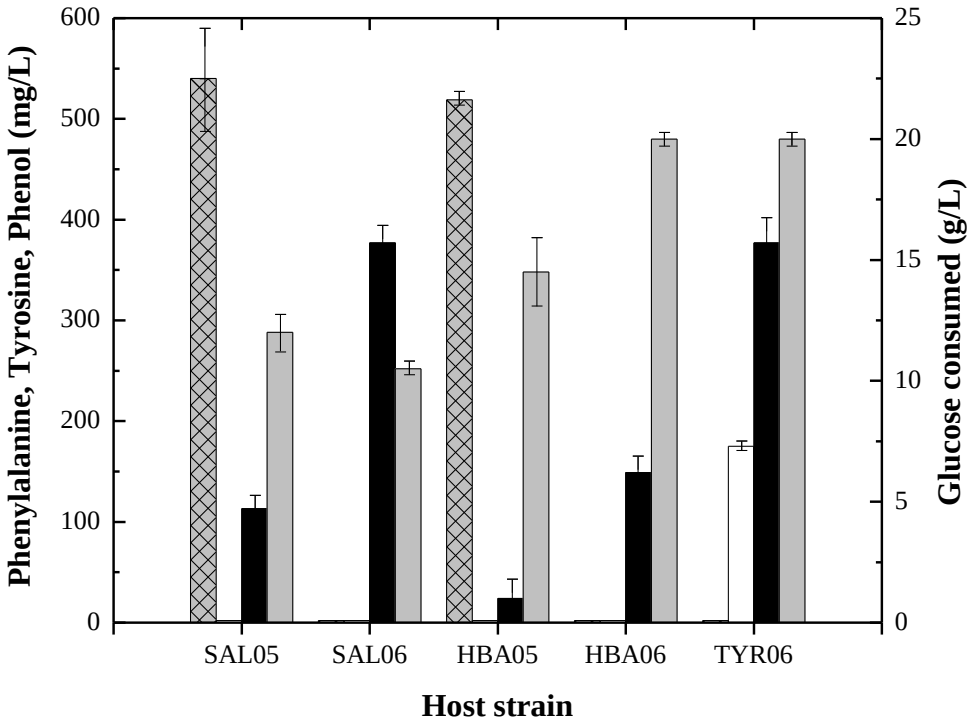


Figure 2.6 Alternative Host Selection for Phenol Production.Effects of employing alternative host strains to reduce the metabolic burden associated with initial, multi-plasmid expression strategies. Residual phenylalanine (hashed), residual tyrosine (white), phenol produced (black), and glucose consumed (gray) following 72 hours of culturing in MM1 minimal media supplemented with 20 g/L glucose. Error bars represents one standard deviation from triplicate experiments.

phenylalanine (as much as 540 mg/L) were now also co-produced. Similar to tyrosine, phenylalanine biosynthesis also directly competes for chorismate and should be disrupted. Accordingly, with *pheA*^{fbr} deleted from its chromosome, NST74 Δ*pheA* was next investigated as a phenol production host. In the case of Pathway 2, phenol production by SAL06 reached peak titers of 377 ± 14 mg/L at a glucose yield of 35.7 ± 0.8 mg/g (Figure 2.6). Compared with SAL01, this output represents a 246 and 561% increase in titer and yield, respectively. Similarly, in the case of Pathway 1, with NST74 Δ*pheA* as the host background, a modest improvement in phenol production was also realized, with achievable titers by TYR06 reaching up to 377 ± 23

mg/L at a glucose yield of 18.7 ± 0.7 mg/g (Figure 2.6). At this output, 867 and 835% increases in phenol titer and yield were achieved by strain TYR06 relative to TYR01. As before, however, inherent inefficiencies associated with this TPL-dependent pathway limited net precursor flux to phenol, with residual tyrosine remaining in the culture at up to 175 ± 5 mg/L (Figure 2.6). In contrast, and as with all other strategies investigated herein, no observable improvements in phenol production were demonstrated in the case of Pathway 3, with final titers by HBA06 reaching just 149 ± 12 mg/L Figure 2.6).

2.3.5. Comparing Relative Performance of Established and Novel Phenol Biosynthesis Pathways

Two novel phenol biosynthesis pathways were successfully constructed and evaluated in terms of their phenol production potential relative to the established, TPL-dependent route. Ultimately, the highest achievable phenol production via the tyrosine-, salicylate-, and *p*- hydroxybenzoate-derived pathways (Pathways 1, 2, and 3, respectively) was 377 ± 23, 377 ± 14, 259 ± 31 mg/L, respectively, with corresponding glucose yields of and 18.7 ± 0.7, 35.7 ± 0.8, and 12.9 ± 1.1 mg/g. Meanwhile, whereas phenol titers were comparable between Pathways 1 and 2, glucose yields were nearly 2-fold higher in the latter case. In contrast, not only were glucose yields significantly lower in the case of Pathway 3, but it was further found that all strain engineering strategies applied in said case were ultimately and wholly unsuccessful (production remained highest when using the original, wild-type host background; HBA01). While the underlying and associated mechanisms for now remain unknown, it is clear that there are a number of confounding and likely complex factors responsible for the reduced efficacy of this pathway.

Phenol production by the strains developed in this study, including by all three pathways, also compares well with prior reports of microbes engineered to produce phenol by just Pathway 1 (a maximum titer of 141 mg/L and a glucose yield of 34 mg/g by *P. putida* and a maximum titer of 419 mg/L and a glucose yield of 2.6 mg/g by *E. coli*) (Kim et al. 2014; Wierckx et al. 2005). It should be noted, however, that phenol titers achieved by Kim *et al.*, although slightly higher than those demonstrated here, relied upon the use of glucose minimal media supplemented with 3 g/L yeast extract. Thus, the resultant boost in phenol titer and yield realized in said study may have ultimately been due to the more abundant availability of nutrients provided by the yeast extract. In contrast, in the present study, as only 0.1 g/L phenylalanine was supplemented to accommodate for a single phenylalanine auxotrophy, phenol was produced solely from glucose. Additional studies employing the same, more complex medium composition would be required to further understand the importance of such differences. Meanwhile, it should also be noted that analogous host backgrounds and culture conditions were selected in the present study in order to enable a fair, head-to-head comparison and evaluation of the relative merits and potential pitfalls associated with Pathways 1-3. Overall, however, it is expected that greater overall phenol production in all three cases should in fact be possible through individual optimization of the chosen host strain and/or culture condition employed.

Whereas a pathway to phenol via salicylate (Pathway 2) has not before been reported, salicylate has served as a key intermediate in support of the synthesis of other target chemicals of interest. For example, Lin et al. recently reported the engineering of a novel pathway to muconic acid in which the intermediate salicylate is converted instead to catechol via salicylate monooxygenase (SMO) before catechol is then transformed to muconic acid by catechol 1,2-dioxygenase (CDO) (Lin et al. 2014). Using a pathway that was likewise composed of *entC* and *pchB*, salicylate biosynthesis

was first demonstrated as a terminal product, reaching titers as high as 1.2 g/L (8.7 mM). While this output represents a molar flux to salicylate that is ~2.2-fold greater than was ultimately demonstrated here (a maximum of 4.0 mM phenol produced from salicylate), it should be noted that said study relied upon the alternative use of a mixture of (in g/L) glycerol (10), glucose (2.5), and yeast extract (1) as substrates, in contrast to the sole use of glucose here. As a non-phosphotransferase system (PTS) substrate, glycerol conserves phosphoenolpyruvate (PEP) levels in the cell and accordingly supports the enhanced synthesis of products derived from the shikimic acid pathway (Gottlieb et al. 2014). Moreover, over-expression of native *ppsA* and *tktA* was also performed, strategies demonstrated to further promote the production PEP and erythrose 4-phosphate (E4P), respectively, the two key precursors to the shikimic acid pathway. Accordingly, it is similarly expected that enhanced phenol production via salicylate (Pathway 2) could likewise be achieved by incorporating such strategies, as will be investigated in future works.

Meanwhile, concurrent to the present study, Miao et al. also recently reported on the successful engineering of a pathway for phenol biosynthesis in *E. coli* via the intermediate *p*- hydroxybenzoate (Pathway 3, Figure 2.1) (Miao et al. 2015a). In their study, however, rather than *kpdBCD*, *ycl* from *E. coli* W was overexpressed to convert *p*-hydroxybenzoate to phenol – an enzyme they ultimately found to be rate limiting in the pathway. Similar to the present study, native *ubiC* and $aroG^{fbr}$ were also co-expressed to complete the pathway and enhance precursor supply, respectively. In the end, maximum phenol titers achieved in shake flasks reached 355 mg/L (3.8 mM), but only when cultures were directly supplemented with 138 mg/L (1 mM) *p*-hydroxybenzoate (the pathway intermediate). More recently, while this manuscript was under revision, Ren et al. also reported on the construction of an alternative phenol biosynthesis pathway, in this case via salicylate (i.e., Pathway 2, Figure 2.1)

(Ren et al. 2015). In this case, as in the present study, each of *entC*, *pchB*, and *SDC* were likewise overexpressed in *E. coli* resulting in the production of 406 mg/L phenol, or up to 472 mg/L after further optimization of the medium formulation; output levels that agree well with those reported herein.

Regardless of the chosen pathway, product toxicity is a common concern associated with the production of aromatic chemicals (McKenna and Nielsen 2011; McKenna et al. 2013), including phenol (Wierckx et al. 2005). Past reports have shown, for example, that *E. coli* growth is completely inhibited in the presence of 1.75 g/L phenol, with overall reductions in growth also observed at lower levels (Kim et al. 2014). Based on this, end product toxicity was considered to be an unlikely factor with respect to influencing overall host fitness under the conditions examined. Nevertheless, as others have found the use of *in situ* solvent extraction to remove phenol from cultures as it is produced as an effective strategy for improving production (Miao et al. 2015a), it may prove analogously effective in the case of strains developed here.

2.4. Conclusions

Two non-natural and novel pathways for phenol biosynthesis were engineered and comparatively evaluated to understand their inherent advantages/limitations and to assess their performance relative to the established route. Under otherwise analogous conditions, the highest achievable phenol production via the salicylate-, *p*-hydroxybenzoate-, and established tyrosine-derived pathway was 377 ± 14, 259 ± 31, and 377 ± 23 mg/L, respectively, with corresponding glucose yields of 35.7 ± 0.8, 12.9 ± 1.1, and 18.7 ± 0.7 mg/g. Despite strain engineering efforts resulting in increased overall phenol production via the tyrosine derived pathway, TPL equilibrium limitations continue to limit the overall pathway efficiency. Meanwhile, in contrast to

the other routes, phenol production via the *p*-hydroxybenzoate derived pathway was highest in the original wild-type host, with all efforts to increase overall pathway flux unsuccessful. While this suggests that other complex regulatory factors may also be specifically associated with this pathway and/or its intermediates, further investigation will likely be required for these to be elucidated and ultimately overcome. Lastly, it should also be noted that while equivalent background strains and culture conditions were investigated to provide a fair pathway comparison, alternative strain selection and optimization of culture conditions relative to each individual pathway presented herein may ultimately result in higher achievable phenol titers.

3. ENGINEERING A NETWROK OF PATHWAYS FOR THE NOVEL BIOSYNTHESIS OF MUCONIC ACID

Abstract

Muconic acid (MA) is promising platform biochemical and precursor to adipic acid, which can be used to synthesize various plastics and polymers. Multiple pathways for *de novo* MA biosynthesis from glucose have previously been reported, each incorporating unique enzyme chemistries and/or stemming from different shikimic acid pathway precursors. Here, three novel routes to MA were systematically engineered in *Escherichia coli* by linking a series of pathways recently constructed for the biosynthesis of phenol with select steps associated with its subsequent, partial aerobic degradation. In addition, a fourth route to MA was also engineered, in this case from endogenous chorismate and proceeding via *p*-hydroxybenzoate, protocatechuate, and catechol. Said novel pathways were comparatively evaluated against each other and the established, 3-dehydroshikimate-derived route, with latter supporting the highest MA titers and yields at 819 mg/L and 40.9 mg/g, respectively, as supported by both model simulations and experimental studies.

Select portions of this chapter have been submitted as:

Thompson, B., Pugh, S., Machas, M., Nielsen, D.R. Exploring Novel Pathways and Synthetic Metabolic Funneling to Enhance Muconic Acid Biosynthesis. *Metab Eng* (2017).

3.1. Introduction

In recent years, muconic acid (MA) has continued to emerge as a diacid bioproduct of industrial interest, most notably due to its role as a precursor to adipic acid – a platform chemical used for the synthesis of various plastics and polymers (e.g., Nylon 6,6) (Adkins et al. 2012; Deng et al. 2016). Through metabolic pathway engineering, numerous studies have reported the engineering of microbes capable of producing MA as the focal product from renewable feedstocks (Curran et al. 2013; Draths and Frost 1994; Johnson et al. 2016; Lin et al. 2014; Niu et al. 2002; Sengupta et al. 2015; Sun et al. 2013; Sun et al. 2014; Weber et al. 2012; Zhang et al. 2015a; Zhang et al. 2015c). Draths and Frost were first to report MA biosynthesis from glucose in *Escherichia coli*, following construction of a two-step pathway stemming from endogenous 3-dehydroshikimate (3DHS), a key intermediate in the shikimic acid pathway (Draths and Frost 1994). Said pathway, which has since also been functionally re-constructed in *Saccharomyces cerevisiae* (Curran et al. 2013; Weber et al. 2012), proceeds through the intermediates protocatechuate (PCA) and catechol via 3-DHS dehydratase, PCA decarboxylase, and catechol 1,2-dioxygenase (hereafter referred to as '3DHS-derived'). Achieving significant MA production via this pathway has typically required deletion of *aroE* (encoding shikimate dehydrogenase, an essential gene in minimal media). This mutation results in auxotrophies for the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp), as well as the growth essential vitamins *p*-aminobenzoate (pAB), *p*-hydroxybenzoate (pHBA), and 2,3-dihydroxybenzoate (2,3-DHB) – all of which are derived from chorismate as their last common precursor. Thus, although high MA titers and yields have been achieved via the '3DHS-derived' pathway when expressed in *E. coli ΔaroE* backgrounds (2.4 g/L at 0.24 g/g-glucose, respectively, in shake flasks (Draths and Frost 1994), and 36.8 g/L at 0.17 g/g-glucose in a fed-batch bioreactor (Niu et al.

2002)), doing so required each of the above six nutrients to first be supplemented into the minimal salts media – an expensive and poorly scalable practice.

To address this limitation, multiple alternative MA biosynthesis pathways have recently been engineered, each involving different enzyme chemistries and, most notably, stemming from precursors further downstream in the shikimic acid pathway, thereby preserving chorismate biosynthesis. Because native biosynthesis of each of Phe, Tyr, Trp, pAB, pHBA, and 2,3-DHB is tightly regulated downstream of chorismate (i.e., by feedback repression), potential auxotrophies for these species are effectively avoided with only minimal loss of MA precursors. To date, no fewer than four alternatives pathways have so far been proposed and/or developed for MA biosynthesis from glucose (Averesch and Krömer 2014). For example, Sun et al. recently reported a novel MA biosynthesis pathway utilizing anthranilate (intermediate to Trp) as its immediate endogenous precursor (Sun et al. 2013). By said pathway, MA titers of 0.39 g/L were achieved using a mixed glucose/glycerol minimal media (supplemented with Trp to account for deletion of *trpD*) in shake flask cultures. The same group also later engineered another alternative MA pathway, in this case utilizing endogenous 2,3-DHB as its immediate endogenous precursor (Sun et al. 2014). In this case, MA titers of 0.48 g/L were reported using a glucose/glycerol media with yeast extract in shake flasks cultures. Most recently, the same group has engineered a third, alternative MA pathway that instead stems directly from chorismate and proceeds via the key intermediates isochorismate and salicylate (Lin et al. 2014). Said pathway resulted in MA titers reaching 1.45 g/L in shake flask cultures using the same glucose/glycerol media with yeast extract (in this case to account for Phe and Tyr auxotrophies caused by deletion of *pheA* and *tyrA*, respectively, to increase chorismate availability).

In addition to these novel MA biosynthesis pathways reported to date, alternative strategies for catechol and MA biosynthesis were investigated here by: i) constructing a series of modular, phenol-derived catechol and MA pathways by linking three recently engineered phenol biosynthesis pathways (Thompson et al. 2016) with their subsequent aerobic degradation, and; ii) engineering a four-step MA pathway from endogenous chorismate via the intermediates pHBA, PCA via a previously reported novel catechol biosynthesis pathway (Pugh et al. 2014). Both strategies importantly circumvent the auxotrophic limitations experienced via the original '3DHS-derived' route.

3.2. Materials and methods

3.2.1. Strains and Media

All strains used in this study are listed in Tables 3.1 and 3.2. *E. coli* NEB10-beta (New England Biolabs (NEB); Ipswich, MA) was used for all cloning and plasmid maintenance. *E. coli* NST74 (ATCC 31884) was obtained from the American Type Culture Collection (ATCC; Manassas, VA) and served as the parent strain in this study. *E. coli* JW2580-1 was obtained from the Coli Genetic Stock Center (CGSC; New Haven, CT) and served as the genetic source for the *pheA*::Kan^R cassette. *E. coli* BW25113 was obtained from the CGSC and served as the genetic source for *ubiC* and *entC*. *Citrobacter braakii* (ATCC 29063) was obtained from the ATCC and served as the genetic source for *tutA*. *Klebsiella pneumoniae* PZH572 (ATCC 25955) was obtained from the ATCC and served as the genetic source of *kpdBCD* and *aroY*. *Pseudomonas aeruginosa* PAO1 (DSMZ 22644) was obtained from the Leibniz Institute German Collection of Microorganisms and Cell Cultures and served as the genetic source of *pobA*. *P. stutzeri* OX1 (ATCC BAA-172) was obtained from the ATCC and served as the genetic source of *phKLMNOP*. *Pseudomonas putida* KT2440 (ATCC 47054) was

obtained from the ATCC and served as the genetic source for *catA*. *Corynebacterium glutamicum* (ATCC 13032) was obtained from the ATCC and served as the genetic source of *qsuB*.

Seed cultures of *E. coli* strains were cultured in Luria-Bertani (LB) broth at 32ºC and supplemented with 100 mg/L ampicillin and/or 35 mg/L kanamycin, as appropriate. For catechol and MA biosynthesis, shake flasks were cultured at 32ºC in M9M minimal media supplemented with appropriate antibiotics. M9M was composed of the following (in g/L): Na_2HPO_4 (6), KH_2PO_4 (3), NaCl (0.5), NH_4Cl (2), $MgSO_4 \cdot 7H_2O$ (0.493), $CaCl_2 \cdot 2H_2O$ (0.0147), and glucose (20). Trace elements were supplemented as follows (in mg/L): $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (0.37), H_3BO_3 (2.5), $CoCl_2 \cdot 6H_2O$ (0.714), $CuSO_4$ (0.16), $MnCl_2 \cdot 4H_2O$ (1.6), $ZnSO_4 \cdot 7H_2O$ (0.288), $FeCl_3$ (0.05).

3.2.2. Plasmid Construction

All plasmids used and developed in this study are listed in Table 3.1. All genes were PCR amplified with Q5 High-Fidelity DNA Polymerase (NEB) and a BioRad iCycler, per manufacturer protocols. Custom DNA oligonucleotide primers (See Table B1, Appendix B) were synthesized by Integrated DNA Technologies (IDT, Coralville IA). Genomic DNA (gDNA) templates were prepared using the ZR Fungal/Bacterial DNA MiniPrep kit while plasmid DNA was purified using the Zymo Plasmid MiniPrep kit (both Zymo Research, Irvine CA). Amplified linear DNA fragments were purified using the Zymo DNA Clean & Concentrator MiniPrep kit (Zymo Research). Purified linear DNA and plasmid DNA were digested using appropriate restriction endonucleases (NEB) and subsequently gel purified using the Zymoclean Gel DNA Recovery MiniPrep kit (Zymo Research). Purified digested DNA fragments were ligated using T4 DNA Ligase (NEB), per manufacturer protocols. Ligation reactions were transformed into chemically competent *E. coli* NEB10-beta before plating on LB solid agar media supplemented

with appropriate antibiotics for selection. Transformant pools were screened using colony PCR, restriction digest mapping, and finally confirmed by DNA sequencing.

3.2.3. Strain Construction

All strains engineered and used in this study are listed in Tables 3.1 and 3.2. Chromosomal in-frame deletions of *pheA*^{fbr} in *E. coli* NST74 was constructed using a modified version of the Datsenko and Wanner method (Datsenko and Wanner 2000), as previously described (Pugh et al. 2014). The *pheA*:: *FRT-kan*^{R}*-FRT* deletion cassette was PCR amplified from *E. coli* JW2580-1. Chromosomal integration of said cassette and subsequent removal of *kan*^{R} marker was achieved as previously described (Datsenko and Wanner 2000; Pugh et al. 2014), resulting in the individual construction of *E. coli* NST74 Δ*pheA*^{fbr}. For catechol biosynthesis, *E. coli* NST74 Δ*pheA*^{fbr} was co-transformed with the following combinations of plasmids (note: pathway designation provided in parentheses): pY3 and pTutA-pPh (CAT2), pSDC-PchB-EntC and pPh (CAT3), pUbiC-Kpd and pPh (CAT4), and pUbiC-PobA and pAroY (CAT5). For MA biosynthesis, *E. coli* NST74 Δ*pheA* was co-transformed with the following combinations of plasmids: pQsuB-AroY-CatA (MA1), pY3 and pTutA-pPh-CatA (MA2), pSDC-PchB-EntC and pPh-CatA (MA3), pUbiC-Kpd and pPh-CatA (MA4), and pUbiC-PobA and pAroY-CatA (MA5).

3.2.4. Thermodynamic and Elementary Flux Mode Analysis

To compare relative pathway energetics, net changes in Gibbs free energy due to reaction, $\Delta_r G'^{\circ}_{net}$, were calculated using the online eQuilibrator tool (available at http://equilibrator.weizmann.ac.il) at a reference state of 25ºC, pH 7, and ionic strength of 0.1 M. Elementary flux modes (EFMs) were computed in MATLAB R2014b (MathWorks, Natick MA) using EFMTool 4.7.1 (Terzer and Stelling 2008). An *E. coli*

stoichiometric network originally employed to compare different phenol biosynthesis pathways (Thompson et al. 2016) and originally adapted from Averesch and Krömer (2014), was used to compare relative maximum yields from the various MA pathways and strains of interest in this study (see Appendix B).

3.2.5. Assaying Phenol Hydroxylase and Catechol 1,2-Dioxygenase Activity Using Whole Resting Cells

Recombinant activities of phenol hydroxylase and catechol 1,2-dioxygenase were assayed in *E. coli* BW25113 following its transformation with pPh or pPh-CatA. Overnight seed cultures were used to inoculate (1% vol.) 50 mL of LB supplemented with 20 g/L glucose and 35 mg/L kanamycin in 250 mL shake flasks. In addition, the effect of $Fe(NH_4)_2(SO_4)_2$ was also examined via its omission or inclusion at 100 µM. Shake flasks were induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.4 mM upon reaching an optical density at 600 nm (OD_{600}) of 0.7. Culturing continued overnight (~12 h) at 32ºC before cells were then collected by centrifugation at 3,000 x *g* for 5 min. Cell pellets were rinsed twice with pH 6.8 phosphate buffered saline (PBS) before being re-suspended to a final OD_{600} of 4 in PBS supplemented with 0.2 g/L glucose and 1 mM phenol in a 250 mL shake flask. Cultures were subsequently incubated at 32ºC with shaking at 200 RPM for up to 12 hours. Samples were periodically drawn for metabolite quantification via HPLC, as described below.

Table 3.1 Strains and Plasmids Used for MA Production

Strain	Description	Source
E. coli NEB10-beta	*Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)*	NEB
E. coli BW25113	Source of *ubiC* and *entC*	CGSC
E. coli JW2580-1	Source of *pheA*::Kan^R	CGSC
E. coli NST74	*aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*	ATCC
E. coli NST74 Δ*pheA*	*pheA* chromosomal deletion in *E. coli* NST74	Pugh et al. (2014)
C. glutamicum	Source of *qsuB*	ATCC
C. braakii	Source of *tutA*	ATCC
P. aeruginosa PAO1	Source of *pobA* and *pchB*	DSMZ
E. coli BW25113	Source of *ubiC* and *entC*	CGSC
K. pneumoniae PZH572	Source of *kpdBCD* and *aroY*	ATCC
P. stutzeri OX1	Source of *phKLMNOP*	ATCC
P. putida KT2440	Source of *catA*	ATCC
Plasmid	**Description**	**Source**
pTrc99A	pBR322 ori, Amp^r, *lacIq*, P_{trc}	Prather Lab, MIT
pTrcCOLAK	ColA ori, Kan^r, *lacIq*, P_{trc}	McKenna et al. (2013)
pY3	p15A, Amp^r, *lacI*, $P_{lac-UV5}$-*tyrB*-*tyrA*^{fbr}-*aroC* T1-P_{trc}-*aroA*-*aroL*	Juminaga et al. (2012)
pKD46	repA101(ts) and R101 ori, Amp^r, *araC, araBp*	CGSC
pCP20	FLP, ts-rep, [cI857](lambda)(ts), Amp^r	CGSC
pTutA-Ph	*tutA* of *C. braakii* inserted to pPh	This study
pTutA-Ph-CatA	*tutA* of *C. braakii* inserted to pPh-CatA	This study
pSDC-PchB-EntC	SDC of *T. moniliforme, pchB* of *P. aeruginosa* PAO1, and *entC* of *E. coli* BW25113 inserted to pTrc99A	This study
pUbiC-Kpd	*ubiC* of *E. coli* BW25113 and *kpdBCD* of *K. pneumoniae* PZH572 inserted to pTrc99A	This study
pPh	*phKLMNOP* of *P. stutzeri* OX1 inserted to pTrcCOLAK	This study
pPh-CatA	*catA* of *P. putida* KT2440 inserted to pPh	This study
pUbiC-PobA	*ubiC* of *E. coli* BW25113 and *pobA* of *P. aeruginosa* PAO1 inserted to pTrc99A	Pugh et al. (2014)
pAroY	*aroY* of *K. pneumoniae* PHZ572 inserted to pTrcCOLAK	Pugh et al. (2014)
pAroY-CatA	*catA* of *P. putida* KT2440 inserted to pAroY	This study
pQsuB-AroY-CatA	*qsuB* of *C. glutamicum* inserted to pTrc99A	This study

Table 3.2 Strains Engineered for Catechol and MA Biosynthesis

Pathway	Plasmid Combination	Source
CAT2	pY3, pTutA-Ph	This study
CAT3	pUbiC-Kpd, pPh	This study
CAT4	pSDC-PchB-EntC, pPh	This study
CAT5	pUbiC-PobA, pAroY	Pugh et al. (2014)
MA1	pQsuB-AroY-CatA	This study
MA2	pY3, pTutA-Ph-CatA	This study
MA3	pUbiC-Kpd, pPh-CatA	This study
MA4	pSDC-PchB-EntC, pPh-CatA	This study
MA5	pUbiC-PobA, pAroY-CatA	This study

3.2.6. Catechol and MA Production from Glucose

To investigate the bioproduction of catechol and MA, overnight seed cultures were first prepared in LB and supplemented with 100 mg/L ampicillin and/or 35 mg/L kanamycin and used to inoculate (1% vol) 50 mL of M9M minimal media supplemented with 20 g/L glucose in 250 mL shake flasks (note: the medium was additionally supplemented with 0.1 g/L Phe when using Δ*pheA*-derived host strains). Shake flask cultures were incubated at 32ºC with shaking at 200 RPM until reaching OD_{600} ~0.7, at which point IPTG induction was performed at a final concentration of 0.4 mM. Cultures were further incubated for a total of up to 120 h, or until significant sugar consumption was no longer detected. Periodically, samples were drawn to measure cell growth (as OD_{600}) sugar and metabolite levels by HPLC analysis, as described below. Prior to centrifugation (i.e., to pellet and remove cells), samples for MA analysis were first diluted 1:10 with methanol, while samples for Tyr analysis were diluted 1:10 with 1 N HCl and incubated at 55°C for 30 min. All samples were then centrifuged at 11,000 x *g* for 5 min before transferring the supernatant to a glass HPLC vial.

3.2.7. HPLC Metabolite Analysis

Metabolite analysis was performed using a Hewlett Packard 1100 series HPLC system. Separation of Phe, *p*HBA, PCA, salicylate, phenol, catechol, and MA, was achieved using a reverse-phase Hypersil GOLD aQ C18 column (3 mm x 250 mm; Thermo Fisher, Waltham, MA, USA) operated at 45°C with an isocratic 0.8 mL/min mobile phase consisting of 85% (vol.) 5 mM H_2SO_4 and 15% (vol.) acetonitrile. The eluent was monitored using a diode array detector (DAD) set at 215 nm for Phe, salicylate, PCA, and catechol, 260 nm for *p*HBA, and 275 nm for phenol and MA. Separation of Tyr was also achieved on Hypersil GOLD aQ C18 column, in this case maintained at 30°C while using a mobile phase consisting of water (A) and methanol plus 0.1% (vol.) formic acid (B) at a constant flow rate of 0.2 mL/min and the following concentration gradient (all by vol.): 5% B from 0 to 8 min, 5% to 40% B from 8 to 13 min, 40% B from 13 to 16 min, 40 to 5% B from 16 to 21 min, and 5% B from 21 to 31 min. The eluent was monitored using a DAD at 215 nm. Glucose and acetate separation was achieved using an Aminex HPX-87H column (BioRAD, Hercules, CA) operated at 35ºC and detected using a refractive index detector (RID). The column was eluted with 5 mM H_2SO_4 at a constant flow rate of 0.55 mL/min. In all cases, external standards were prepared and used to provide calibrations for concentration determination.

Mass spectrometry analysis was performed using a Dionex Ultimate 3000 HPLC system (Bruker Daltonics, Billerica, MA, USA) consisting of a HPG-3400 M pump, WPS 3000 TB autosampler fitted with a 5 µL sample loop, and a FLM-3100B thermostatted column compartment. A Hypersil GOLD aQ C18 column (3 mm x 250 mm; Thermo Fisher, Waltham, MA, USA) was operated at 25°C with an isocratic 0.2 mL/min mobile phase consisting of 85% (vol.) 5 mM formic acid and 15% (vol.) acetonitrile. Metabolites were detected as negative ions using a Bruker MicrOTOF-Q mass

spectrometer with the following settings: Capillary voltage, +4000; end plate offset, -500V; nebulizer gas pressure, 2.0 bar; dry gas flow, 8 L/min; dry gas temperature, 210° C; Funnels 1 and 2 radio frequency (RF) settings, 200 Vpp, Hexapole RF setting, 150 Vpp; Collision Cell RF setting, 100 Vpp; Quadrupole low mass setting, 55 m/z; Transfer time, 100 µs; Pre Pulse Storage, 7 µs. A Peak Scientific NM30LA nitrogen generator (Peak Scientific Inc., Billerica, MA, USA) supplied nitrogen for the drying and nebulizer gases.

3.3. Results and Discussion

3.3.1. Novel Pathway Identification and Theoretical Comparison

In addition to the original, '3DHS-derived' pathway (steps A, B, C in Figure 3.1; hereafter referred to as MA1) (Draths and Frost 1994), as discussed above, several alternative pathway options for MA biosynthesis have since also been developed. Although proceeding from different endogenous metabolites and/or involving alternative enzyme chemistries, catechol always serves as the immediate MA precursor. However, in addition to serving as the ubiquitous MA precursor, catechol is also the first intermediate associated with aerobic phenol catabolism – a degradation pathway common to many soil microbes, including various *Pseudomonas* sp., for example (van Schie and Young 2000). Accordingly, it was hypothesized that, by linking engineered phenol production with its partial, aerobic degradation, additional new routes to MA could ultimately also be engineered. More specifically, as illustrated in Figure 3.1, this could be achieved by further co-expressing phenol hydroxylase (PH) and catechol 1,2-dioxygnease (CDO) (steps K, C; Figure 3.1) in a phenol-producing background. Whereas engineered phenol production has traditionally involved expression of tyrosine phenol lyase (TPL; step J) in a Tyr overproducing host (Kim et al. 2014; Wierckx et al. 2005), two alternative phenol biosynthesis pathways from

chorismate were recently reported, involving either isochorismate synthase, isochorismate pyruvate lyase, and salicylate decarboxylase (steps G, H, I) or chorismate lyase and pHBA decarboxylase (steps D, F) (Thompson et al. 2016). Accordingly, three novel, phenol-dependent MA biosynthesis pathways were proposed, hereafter referred to as MA2 (steps J, K, C), MA3 (steps G, H, I, K, C), and MA4 (steps D, F, K, C). Meanwhile, in separate works, a novel pathway for catechol biosynthesis was previously reported that proceeds from chorismate through pHBA and PCA via chorismate lyase, pHBA hydroxylase, and PCA decarboxylase (steps D, E, B) (Pugh et al. 2014). Accordingly, further extension of this pathway to MA was also proposed, in this case by co-expressing CDO, resulting in pathway MA5 (steps D, E, B, C).

Unlike MA1, the four alternative pathways proposed herein (MA2-5) importantly stem from chorismate or one of its downstream metabolites, and thereby offer improved host compatibility by preserving native flux through the shikimic acid pathway. In addition, relative to MA1, each of MA2-5 was also found to further benefit from an increased net thermodynamic driving force (as much as 104% greater). More specifically, when evaluated from the last common precursor (i.e., 3DHS), the net change in Gibbs free energy due to reaction ($\Delta_r G'°_{net}$) was predicted to be -1037.4 kJ/mol for each of MA3-5 and -1007.6 kJ/mol for MA2, compared to just -508.5 kJ/mol for MA1 (Figure 3.2). In contrast, as determined via elementary mode analysis (Figures B1a-e, Appendix B), MA1 supports the highest product yield, with the predicted maximum theoretical yields of MA from glucose ($Y_{P/S,max}$) being ~18% lower (or ~15% lower with growth, $Y_{P/S,max+growth}$; Figure 3.2) for each of MA2-5. In all cases, said reduction is due to the additional 1 NADPH (by AroE), 1 ATP (by AroL), and 1 phosphoenolpyruvate (PEP; by AroA) consumed while converting 3DHS to chorismate. Meanwhile, both MA3 and MA4 further require an additional 1 NADH (i.e., by PH; note: this demand is balanced in MA2 by the generation of 1 NADH by TyrA), whereas MA5

requires an additional 1 NADPH (by pHBA hydroxylase). Other previously-reported chorismate-derived MA pathways (i.e., derived from anthranilate or from salicylate, via an alternative mechanism (Lin et al. 2014; Sun et al. 2013)) suffer the same limitations with respect to predicated maximum MA yields from glucose, in this case also being about 17-20% lower than MA1 (or 18-19% lower with growth; note: the anthranilate-derived pathway offers the lowest yield due to a requirement for an additional 1 ATP) (Averesch and Krömer 2014). Therefore, despite incorporating more favorable enzyme chemistries and the potential for improved host compatibility, potential improvements associated with the proposed and other MA pathways come at the cost of lower maximum theoretical yields.

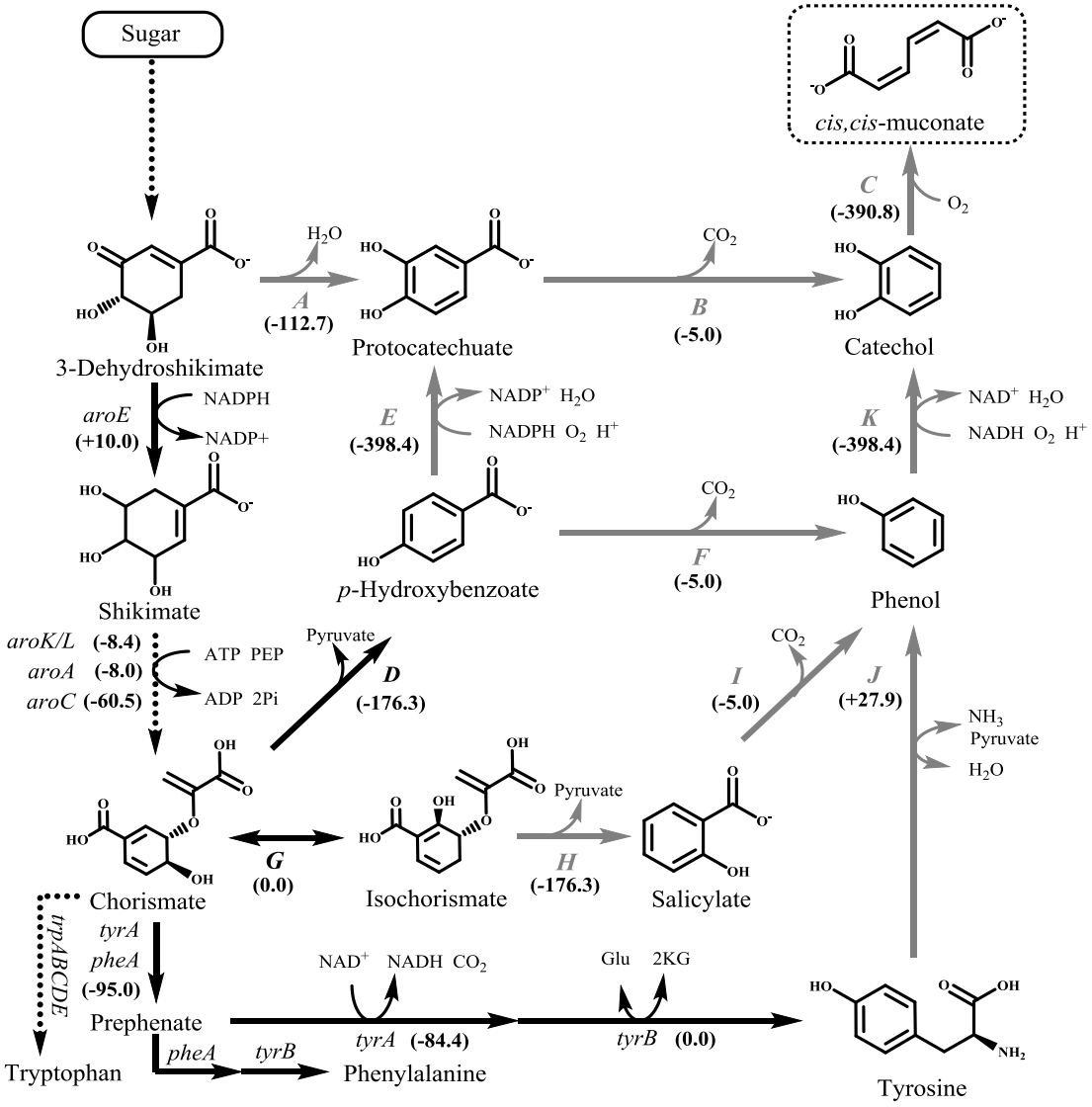
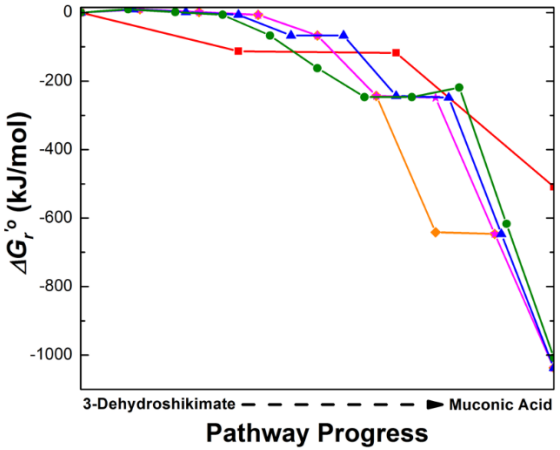


Figure 3.1 Five Non-natural Pathways Engineered for Muconic Acid Biosynthesis Black arrows represent native *E. coli* enzymatic steps; gray arrows represent heterologous steps; dotted arrows represent multiple enzymatic steps. $\Delta_r G'°$ is the change in Gibbs free energy due to reaction as determined using the online tool eQuilibrator (*http://equilibrator.weizmann.ac.il*) at a reference state of 25°C, pH 7, and ionic strength of 0.1 M. A: 3-dehydroshikimate dehydratase, B: protocatechuate decarboxylase, C: catechol 1,2-dioxygenase, D: chorismate pyruvate lyase, E: *p*-hydroxybenzoate hydroxylase, F: *p*-hydroxybenzoate decarboxylase, G: isochorismate synthase, H: isochorismate pyruvate lyase, I: salicylate decarboxylase, J: tyrosine phenol lyase, K: phenol hydroxylase.



MA1	3DHS → PCA → Catechol → MA
MA2	Tyr → Phenol → Catechol → MA
MA3	Chorismate → Isochorismate → Salicylate → Phenol → Catechol → MA
MA4	Chorismate → pHBA → Phenol → Catechol → MA
MA5	Chorismate → pHBA → PCA → Catechol → MA

	$\Delta G_r'^{0}$ (kJ/mol)	$Y_{X/S,max}$	$Y_{P/S,max}$	$Y_{P/S,max\ +\ growth}$	*EM Count*
MA1	-508.5	0.7086	0.8571	0.7945	468,715
MA2	-1007.6	0.7086	0.7013	0.6609	676,940
MA3	-1037.4	0.7086	0.7059	0.6777	506,231
MA4	-1037.4	0.7086	0.7059	0.6777	506,231
MA5	-1037.4	0.7086	0.7059	0.6777	506,231

Figure 3.2 Comparing the Change in Gibbs Free Energy of Reaction ($\Delta_r G'°$) as a function of pathway progress from 3-DHS (the last common endogenous precursor) to MA for each of the five pathways studied (MA1: red squares, MA2: green circles, MA3: blue triangles, MA4: pink stars, MA5: orange diamonds). Also compared are the predicted maximum MA yields ($Y_{P/S}$), MA yields with growth ($Y_{P/S+growth}$), and biomass yields ($Y_{X/S}$) from glucose, as determined by elementary mode (EM) analysis, and the corresponding EM Counts associated with each simulation. PCA: protocatechuate, Tyr: tyrosine, pHBA: p-hydroxybenzoate.

3.3.2. Screening for and Characterizing Phenol Hydroxylase Activity in *E. coli*

Effective enzyme candidates for most steps associated with each of the proposed pathways have been reported and/or characterized by others, or have appeared as part of our own previous works (Lin et al. 2014; Pugh et al. 2014; Thompson et al. 2016). Phenol-dependent MA biosynthesis (i.e. via pathways MA2-4) represents a new approach, however, and first required identification of a candidate gene whose expression would confer recombinant PH activity in *E. coli*. PH activity has been identified as part of bacterial multicomponent monooxygenases (BMMs). BMMs represent a broad class of enzymes capable of using molecular oxygen to hydroxylate various hydrocarbon species, and have been identified to function in numerous microbes as the initial activating steps involved in degrading benzene, toluene, and xylene (Jindrová et al. 2002; Notomista et al. 2003; Sridevi et al. 2012). Notable examples include toluene/o-xylene monooxygenase (ToMo, encoded by *touABCDEF*) and phenol hydroxylase (PH, encoded by *phKLMNOP*), both from *Pseudomonas stutzeri* OX1, which together function to sequentially catalyze the first two steps in benzene degradation (i.e., via phenol then catechol) (Cafaro et al. 2004; Cafaro et al. 2002; Tinberg et al. 2011). As promiscuous enzymes, previous *in vitro* studies have shown that both *touABCDEF* and *phKLMNOP* display the desired PH activity; however, *phKLMNOP* expectedly exhibits more favorable activity towards phenol (K_m=0.6 μM, k_{cat}=1.02 s^{-1}), its native substrate, than does *touABCDEF* (K_m=2.18 μM, k_{cat}=1.00 s^{-1}) (Cafaro et al. 2004). Accordingly, *phKLMNOP* was selected as the initial candidate for establishing recombinant PH activity (step K; Figure 3.1) in *E. coli*. Meanwhile, whereas CDO activity (step C) has been identified in a variety of aromatic-degrading microorganisms (Bouwer and Zehnder 1993; Cao et al. 2008; Kukor et al. 1988), CatA from *P. putida* displays high recombinant activity

(10.10 ± 0.35 μmol/min/mg protein) (Sun et al. 2013), and thus was accordingly used in each of MA1-5.

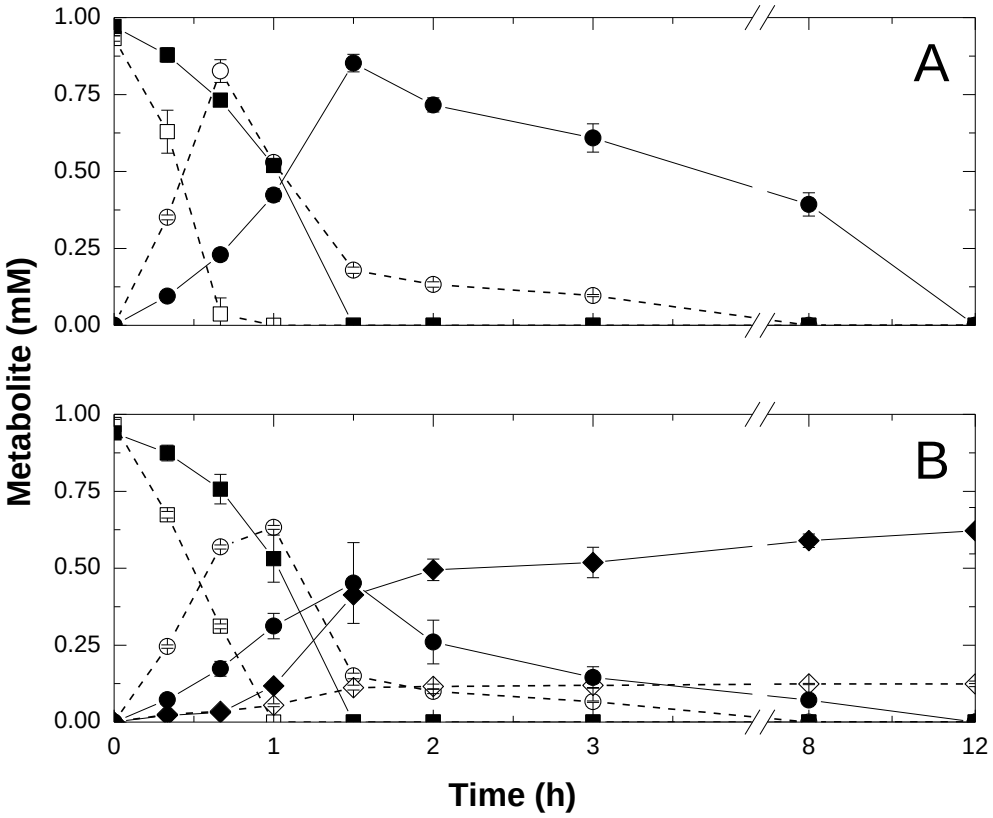


Figure 3.3 Screening Candidate Phenol Hydroxylase ActivityEnzyme was assayed using whole resting cells with (dashed lines, open symbols) and without (solid lines, filled symbols) 100 μM $Fe(NH_4)_2(SO_4)_2$ supplementation. A) Conversion of phenol (squares) to catechol (circles) by *E. coli* BW25113 pPh. B) Conversion of phenol (squares) to catechol (circles) or muconic acid (diamonds) by *E. coli* BW25113 pPh-CatA. Error bars represent one standard deviation from triplicate experiments.

Recombinant PH activity was confirmed *in vivo* via whole resting cell assays employing *E. coli* BW25113 pPh. Because PH possesses carboxylate-bridged diiron catalytic centers in each of its N subunits, as well as a [2Fe-2S] cluster in the P subunit (Cafaro et al. 2004; Tinberg et al. 2011), the medium was first supplemented with 100 μM $Fe(NH_4)_2(SO_4)_2$. Under these conditions, 1 mM exogenous phenol was rapidly

converted to catechol, achieving a maximum specific rate of 0.991 mmol g^{-1} h^{-1} (Figure 3.3A). However, after reaching a maximum level of 0.83 ± 0.04 mM after 40 min, catechol levels then gradually decreased over time, ultimately being undetected after 8 h. It was initially hypothesized that the disappearance of catechol was a result of dimer formation, as said phenomena have been reported to occur for various catechols in the presence of excess Fe(III) and under mildly acidic conditions, a process that plays an important role in adhesive formation in the mussel byssus (Fullenkamp et al. 2014). To test this hypothesis, and determine if such an undesirable side reaction could perhaps be avoided, the assay was next repeated without $Fe(NH_4)_2(SO_4)_2$ supplementation (Figure 3.3A). Under these conditions, phenol conversion to catechol occurred at a similar yield (0.85 ± 0.03 mM catechol produced from 1 mM phenol after 1.5 h), proceeding, however, at only ~60% of the previous maximum specific rate (just 0.596 mmol g^{-1} h^{-1}). However, even in the absence of Fe(III), catechol was again depleted following its formation, here completely disappearing after 12 h. As an alternative hypothesis, several BMMs have been reported to display promiscuous activity towards numerous, related aromatic species, including the repeated hydroxylation of the same initial substrate molecule (Tao et al. 2004; Vardar and Wood 2005). Thus, it was postulated that the PH encoded by *phKLMNOP* might also be capable of further hydroxylating catechol to 1,2,3-trihydroxybenzene (1,2,3-THB). LC-MS analysis was performed on supernatants prepared from the above resting cell assays, wherein an unknown peak was identified with a molar mass (m/z) of 125.03 ± 0.02 (Figure B2, Appendix B). As the molecular weight of 1,2,3-THB 126.11 g/mol, this strongly suggests that catechol disappearance is occurring due to its further hydroxylation to 1,2,3-THB by PH. Although previously reported for ToMo (Vardar and Wood 2005), to the best of our knowledge, this is the first report of such behavior by *phKLMNOP* from *P. stutzeri* OX1.

Although it was determined that 1,2,3-THB is not inhibitory to *E. coli* (no significant growth defects observed in the presence of as much as 6 g/L 1,2,3-THB; Figure B3, Appendix B), for any pathway incorporating PH (i.e., MA2-4), co-production of 1,2,3-THB will likely compete for available catechol, thereby reducing MA production. Accordingly, it was hypothesized that rapid turnover of catechol to MA by CDO might enable the undesirable production of 1,2,3-THB to be minimized upon implementation of the full pathway. As a preliminary test, the above experiment was repeated using whole resting cells of *E. coli* BW25113 pPh-CatA which co-express CDO (encoded by *catA* from *P. putida* KT2440) together with PH. In this case, as seen in Figure 3B for both with and without $Fe(NH_4)_2(SO_4)_2$ supplementation, 1 mM phenol was converted to MA at maximum final concentrations of 0.62 ± 0.08 and 0.12 ± 0.01 mM, respectively. Thus, although co-expression of PH and CDO enabled the successful transformation of phenol to MA (which, unlike catechol, stably accumulated for the duration of the experiment; Figure 3.3B), overall conversion was low (reaching no greater than 62%) due to the persistent competing formation of 1,2,3-THB by PH.

3.3.3. Investigating Phenol-Dependent Catechol Production in *E. coli*

Three distinct phenol biosynthesis pathways (Thompson et al. 2016) were each next extended to catechol via the additional co-expression of *phKLMNOP*. This resulted in the construction of three novel catechol production pathways: CAT2 (steps J, K), CAT3 (steps G, H, I, K), and CAT4 (steps D, F, K) (Figure 3.1). When expressed in *E. coli* NST74 Δ*pheA* in shake flask cultures, phenol-dependent catechol biosynthesis from glucose was successfully demonstrated in all three cases, the results of which are compared in Table 3.3. For each of CAT2-4, however, significant levels of residual phenol remained at the end of each culture (reaching as high as 149 ± 1 mg/L and, in all cases, surpassing total catechol production), suggesting low PH activity as a

common bottleneck. In addition, similar to the whole resting cell assays, all three pathways further suffered from the competing formation of 1,2,3-THB, with detected levels reaching as much as 96 ± 4 mg/L. In the case of CAT2, catechol titers reached 79 ± 3 mg/L at a glucose yield of 6.2 ± 0.28 mg/g. In addition to phenol, high levels of unconverted Tyr also accumulated in the medium, reaching 342 ± 12 mg/L after 96 h, pointing to TPL as an additional flux limiting step in the pathway. TPL, for example, is known to suffer from both equilibrium limitations (the reaction is reversible, with $\Delta_r G'^{\circ}$ = +27.9 kJ/mol; Figure 3.1) as well as feedback inhibition by phenol (e.g., 94 mg/L phenol reduces TPL activity to only 23% of its maximum (Wierckx et al. 2008)). Among the three pathways, CAT3 enabled the highest catechol titer and yield, reaching 100 ± 2 mg/L and 12.8 ± 0.26 mg/g, respectively; however, residual salicylate also accumulated at 55 ± 2 mg/L in this case. Moreover, glucose utilization remained low among all strains, with overall consumption averaging just ~44% (Table 3.3), the precise reasons for which presently remain unclear. Thus, although functional and enabling phenol-dependent catechol production from glucose for the first time, all three novel catechol pathways (i.e., CAT2-4) suffer from inherent limitations associated with PH, both in terms of low recombinant activity as well as its apparent promiscuity. For comparison, the production of catechol via another, chorismate-derived pathway, referred to here as CAT5 (steps D, E, B; Figure 3.1) (Pugh et al. 2014) was investigated in parallel. When expressed in the same host background and under identical culture conditions, said pathway enabled final catechol titers reaching up to 451 ± 44 mg/L at a glucose yield of 35.0 ± 3.0 mg/g – outputs 4.5- and 2.7-fold greater, respectively, than those achieved by even CAT3 – and did so without significant terminal accumulation of any pathway intermediates.

Table 3.3 Catechol Production via Proposed Novel Pathways. *E. coli* NST74 Δ*pheA* was used as the host background; results reflect 120h of culturing. Error bars represent one standard deviation from triplicate experiment. [a]Pugh et al. (2014).

Pathway	Accumulated Intermediate(s) (mg/L)	Catechol (mg/L)	Glucose Utilization (%)	$Y_{P/S}$ (mg/g)
CAT2	Tyr: 342 ± 12 Phenol: 149 ± 1	79 ± 3	65 ± 1	6.2 ± 0.28
CAT3	Salicylate: 55 ± 2 Phenol: 111 ± 2	100 ± 2	39 ± 1	12.8 ± 0.26
CAT4	Phenol: 125 ± 9	51 ± 10	27 ± 1	9.4 ± 2.1
CAT5[a]	n.d.	451 ± 44	64 ± 1	35.0 ± 3.0

3.3.4. Investigating MA Production via Newly Engineered Pathways

The relative prospects of MA biosynthesis via the three phenol-dependent pathways (MA2-4) and the previously-reported, chorismate-derived catechol pathway (MA5) were investigated via the additional co-expression of *catA* from *P. putida* KT2440 (encoding CDO). *E. coli* NST74 Δ*pheA* was again employed as the initial host background of interest, and the results are compared in Table 3.4. MA titers by MA2, which proceeds through phenol via Tyr, were lowest, reaching just 186 ± 11 mg/L at a glucose yield of 21.0 ± 2.2 mg/g. As was the case for CAT2, this appears to be due at least in part to flux limitations associated with both TPL and PH, as indicated by the terminal accumulation of as much as 220 ± 12 mg/L Tyr and 63 ± 1 mg/L phenol. Like CAT3, MA3 resulted in the highest MA titers and yields among all phenol-derived pathways (i.e., MA2-4), reaching 484 ± 44 mg/L and 46.7 ± 6.0 mg/g, respectively. However, analogous to the case of catechol production, the common reliance upon PH in each of MA2-4 similarly resulted in production of 1,2,3-THB byproduct in each case,

here reaching as high as 232 ± 17 mg/L. Furthermore, similar to CAT2-4, low glucose consumption was also observed for each of MA2-4, again averaging just 44% (Table 3.4). Meanwhile, perhaps expectedly, MA biosynthesis was highest in the case of MA5, reaching 819 ± 44 mg/L at a yield of 40.9 ± 2.2 mg/g. Lastly, to provide a head-to-head comparison, MA1 (i.e., the original '3DHS-derived' pathway) was also constructed and expressed in the same host background. In this case, MA production reached 1586 ± 11 mg/L at a yield of 79.3 ± 0.53 mg/g. Overall, said results are consistent with the above model predictions that found MA1 to be the highest yield pathway (Figure 3.2). Interestingly, and in contrast to the original Draths and Frost study wherein an *E. coli* Δ*aroE* background was used to promote 3DHS availability (resulting in multiple, undesirable auxotrophies), significant MA production via MA1 was demonstrated here without disrupting the shikimic acid pathway. This suggests that, at least under the present conditions, MA1 is capable of effectively competing against native metabolism for at least significant portion of available 3DHS. Furthermore, it should also be noted that although low PCA decarboxylase activity has been implicated as a limiting factor associated with MA production via MA1 (and here, by extension, perhaps also MA5) (Sonoki et al. 2014), as PCA remained undetected throughout any of the present cultures (Table 3.4), such effects do not appear to be limiting for the strains and conditions studied here.

Table 3.4 Novel Muconic Acid Production. *E. coli* NST74 Δ*pheA* was used as the host background; results reflect 120h of culturing. Error bars represent one standard deviation from triplicate experiment.

Pathway	Accumulated Intermediate Metabolite (mg/L)	MA (mg/L)	Acetate (g/L)	Glucose utilization (%)	$Y_{P/S}$ (mg/g)
MA1	n.d.	1586 ± 11	12 ± 0.2	100 ± 1	79.3 ± 0.5
MA2	Tyrosine: 220 ± 12 Phenol: 63 ± 1	186 ± 11	10 ± 1	44 ± 4	21.0 ± 2.2
MA3	n.d.	484 ± 44	7.1 ± 0.1	52 ± 2	46.7 ± 6.0
MA4	n.d.	230 ± 20	7.1 ± 0.1	36 ± 1	31.8 ± 3.6
MA5	n.d.	819 ± 44	12 ± 0.2	100 ± 1	40.9 ± 2.2

3.3.5. Conclusions

While MA production via three distinct and novel, phenol-dependent pathways was successfully demonstrated, the utility of said routes was ultimately limited by their common reliance on phenol hydroxylase – an enzyme with poor recombinant activity and high promiscuity. Instead, chorismate-dependent production of catechol and ultimately MA by a novel pathway proceeding via *p*-hydroxybenzoate proved a more promising alternative, with titers reaching 819 ± 44 mg/L at a glucose yield of 40.9 ± 2.2 mg/g. Although this output lags significantly behind the original 3DHS-derived route, further strain engineering combined with dynamic synthetic flux regulation may ultimately result in higher achievable MA production metrics.

4. EXPLORING A SYNTHETIC METABOLIC FUNNELING APPROACH TO ENHANCE PHENOL AND MUCONIC ACID BIOSYNTHESIS

Abstract

Phenol and muconic acid (MA) are industrially-relevant platform chemicals used for the synthesis of an array of plastics and polymers. Multiple pathways for the renewable phenol and MA biosynthesis have previously been reported, each incorporating unique enzyme chemistries. This unique flexibility with respect to pathway engineering arises in part because MA is also a common intermediate associated with the biodegradation of various aromatic compounds. Further exploiting this principle, a synthetic 'metabolic funnel' was engineered wherein multiple distinct phenol and MA pathways stemming from different endogenous precursors were expressed in parallel to maximize net precursor assimilation and flux towards the final product. Using this novel strategy, maximal phenol and MA titers exceeding 0.5 and 3 g/L were realized, both representing the highest production metrics reported via any single pathway to date.

Select portions of this chapter have been submitted as:

Thompson, B., Pugh, S., Machas, M., Nielsen, D.R. Exploring Novel Pathways and Synthetic Metabolic Funneling to Enhance Muconic Acid Biosynthesis**.** *Metab Eng* (2017).

4.1. Introduction

Phenol and muconic acid (MA) are important building block molecules used in the synthesis of various plastics and polymers of industrial relevance (Adkins et al. 2012; Deng et al. 2016). Although the complete biosynthesis of phenol and MA from glucose has previously been demonstrated, in both cases, the original engineered pathways suffer from notable inherent limitations. Phenol, for example, was synthesized from endogenous tyrosine via expression of heterologous tyrosine phenol lyase (TPL) activity (Wierckx et al. 2008; Wierckx et al. 2005). Although phenol was successfully synthesized with maximal shake flask titers reaching 141 mg/L (Wierckx et al. 2005), TPL reaction reversibility and enzymatic feedback inhibition ultimately resulted in poor pathway efficiency and limited production metrics (Kim et al. 2014; Wierckx et al. 2008). As an alternative approach, two novel pathways were engineered and comparatively evaluated against the original tyrosine-derived pathway, both of which instead stem from endogenous chorismate (Thompson et al. 2016). Using these pathways, titers reached as high as 377 mg/L, although overall yields via any phenol pathway remained less than 10% of the theoretical maximum (Thompson et al. 2016). Meanwhile, MA biosynthesis was originally constructed via a three-step pathway stemming from endogenous 3-dehydroshikimate (3DHS). Although shake flask cultures resulted in the accumulation of 2.4 g/L MA (Draths and Frost 1994), high carbon flux to MA necessitated the creation of multiple nutrient auxotrophies resulting in expensive medium formulations and poor scalability. As an alternative approach, numerous novel pathways to MA have been reported to date, all of which stem from endogenous shikimic acid pathway precursors downstream of 3DHS, importantly preserving chorismate biosynthesis and avoiding the creation of multiple nutrient auxotrophies (Curran et al. 2013; Johnson et al. 2016; Lin et al. 2014; Sengupta et al. 2015; Sun et al. 2013; Sun et al. 2014; Weber et al. 2012; Zhang et al. 2015a;

Zhang et al. 2015b). However, production metrics by each novel route lag significantly behind those achieved via the 3DHS derived pathway.

The unique flexibility by which phenol and MA can be synthesized via multiple different pathways arises in part because they are key intermediates involved in naturally-occurring aerobic degradation pathways used by various microbes to catabolize aromatic chemicals including, for example, those from plant biomasses (i.e. lignin-derived monomers) (Douglas 1996; Herrmann 1995; Sridevi et al. 2012) and petrochemicals (e.g. BTEX) (Jindrová et al. 2002) (Figure 4.1). In addition, catechol, which serves as the ubiquitous MA precursor, is supplied to the β-ketoadipate pathway (which precedes and supplies substrates to the TCA cycle) from an array of different aromatic compounds, including phenol, via a series of so-called 'funneling pathways' (Bouwer and Zehnder 1993; Linger et al. 2014). Drawing from and inspired by these natural mechanisms, the construction of a unique synthetic 'metabolic funnel' (Figure 4.1) was investigated as a strategy capable of enhancing phenol and MA production via the parallel co-expression of distinct yet converging biosynthesis pathways. As an analogous approach, β-carotene biosynthesis can be achieved via two disparate pathways: (i) an optimized mevalonate (MVA) pathway, and (ii) a hybrid methylerythritol phosphate (MEP) pathway (Pérez-Gil and Rodríguez-Concepción 2013). Through simultaneous co-expression of the MVA and MEP pathways, production metrics were increased by as much as 113-fold relative to the single pathway controls (Yang and Guo 2014). Given these results, utilizing a synthetic 'funneling' approach demonstrates how product titers and yields can be improved relative to single pathway controls.

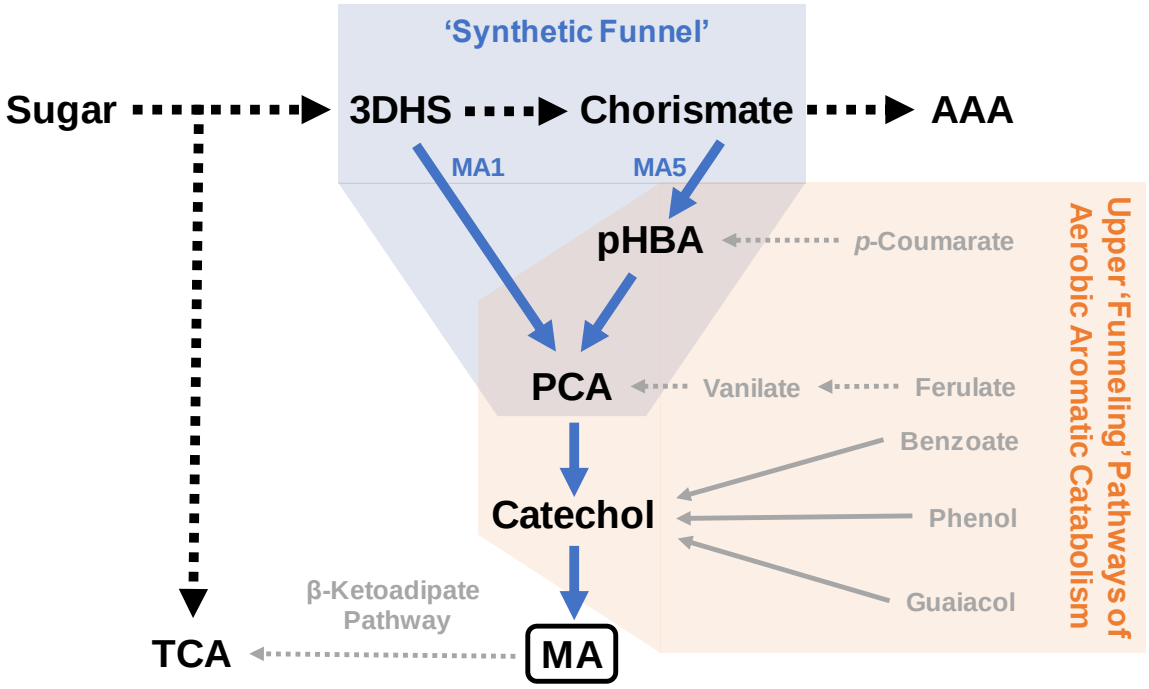


Figure 4.1 Engineering a Synthetic Metabolic 'Funnel'This approach utilizes the simultaneous co-expression of two distinct muconic acid (MA) biosynthesis pathways, stemming from two different precursors (3DHS and chorismate), both derived from the shikimic acid pathway. Also shown are the native mechanisms by which different aromatic degradation pathways also 'funnel' various substrates and intermediates towards MA, in this case for further metabolism via the β-ketoadipate pathway. Black bold dashed arrows represent multiple enzyme reactions native to *E. coli*; blue solid arrows represent individual enzyme steps associated with the pathways engineered in this chapter; gray solid and dashed arrows represent individual and multi-step enzyme processes associated with the aerobic degradation of aromatic chemicals. PCA: protocatechuate; pHBA: *p*-hydroxybenzoate; AAA: aromatic amino acids; TCA: tricarboxylic acid cycle.

4.2. Materials and Methods

4.2.1. Strains and Media

All strains used in this study are listed in Table 4.1. *E. coli* NEB10-beta (New England Biolabs (NEB); Ipswich, MA) was used for all cloning and plasmid maintenance. *E. coli* NST74 (ATCC 31884) was obtained from the American Type Culture Collection (ATCC; Manassas, VA) and served as the parent strain in this study. *E. coli* JW2580-1, JW1843-2, JW1666-3, and JW2410-1 were obtained from the Coli

Genetic Stock Center (CGSC; New Haven, CT) and served as the genetic source for the *pheA*::Kan^R, *pykA*::Kan^R, *pykF*::Kan^R, and *crr*::Kan^R cassettes, respectively. *E. coli* BW25113 was obtained from the CGSC and served as the genetic source for *ubiC* and *entC*. *Citrobacter braakii* (ATCC 29063) was obtained from the ATCC and served as the genetic source for *tutA*. *Klebsiella pneumoniae* PZH572 (ATCC 25955) was obtained from the ATCC and served as the genetic source of *kpdBCD* and *aroY*. *Pseudomonas aeruginosa* PAO1 (DSMZ 22644) was obtained from the Leibniz Institute German Collection of Microorganisms and Cell Cultures and served as the genetic source of *pobA*. *Pseudomonas putida* KT2440 (ATCC 47054) was obtained from the ATCC and served as the genetic source for *catA*. *Corynebacterium glutamicum* (ATCC 13032) was obtained from the ATCC and served as the genetic source of *qsuB*.

Seed cultures of *E. coli* strains were cultured in Luria-Bertani (LB) broth at 32ºC and supplemented with 100 mg/L ampicillin, 35 mg/L kanamycin, and/or 34 mg/L chloramphenicol, as appropriate. For phenol biosynthesis, shake flasks were cultured at 32ºC in MM1 phosphate limited minimal media supplemented with appropriate antibiotics. MM1 was composed of the following (in g/L): $MgSO_4 \cdot 7H_2O$ (0.5), $(NH_4)_2SO_4$ (4.0), MOPS (24.7), KH_2PO_4 (0.3), K_2HPO_4 (0.7), and glucose (20). For MA biosynthesis, shake flasks were cultured at 32ºC in M9M minimal media supplemented with appropriate antibiotics. M9M was composed of the following (in g/L): Na_2HPO_4 (6), KH_2PO_4 (3), NaCl (0.5), NH_4Cl (2), $MgSO_4 \cdot 7H_2O$ (0.493), $CaCl_2 \cdot 2H_2O$ (0.0147), and glucose (20). Trace elements were supplemented in MM1 and M9M as follows (in mg/L): $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (0.37), H_3BO_3 (2.5), $CoCl_2 \cdot 6H_2O$ (0.714), $CuSO_4$ (0.16), $MnCl_2 \cdot 4H_2O$ (1.6), $ZnSO_4 \cdot 7H_2O$ (0.288), $FeCl_3$ (0.05).

Table 4.1 Strains and Plasmids Used for Phenol and MA Production

Strain	Description	Source
E. coli NEB10-beta	Δ*(ara-leu) 7697 araD139 fhuA* Δ*lacX74 galK16 galE15 e14-* φ*80*d*lacZ*Δ*M15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1* Δ*(mrr-hsdRMS-mcrBC)*	NEB
E. coli BW25113	Source of *ubiC* and *entC*	CGSC
E. coli JW2580-1	Source of *pheA*::Kan^R	CGSC
E. coli JW1843-2	Source of *pykA*::Kan^R	CGSC
E. coli JW1666-3	Source of *pykF*::Kan^R	CGSC
E. coli JW2410-1	Source of *crr*::Kan^R	CGSC
E. coli NST74	*aroH367, tyrR366, tna-2, lacY5, aroF394(fbr), malT384, pheA101(fbr), pheO352, aroG397(fbr)*	ATCC
E. coli NST74 Δ*pheA*	*pheA* chromosomal deletion in *E. coli* NST74	Pugh et al. (2014)
E. coli NST74 Δ*pheA ΔpykA ΔpykF Δcrr*	*crr* chromosomal deletion in *E. coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF*	This study
C. glutamicum	Source of *qsuB*	ATCC
C. braakii	Source of *tutA*	ATCC
P. aeruginosa PAO1	Source of *pobA* and *pchB*	DSMZ
E. coli BW25113	Source of *ubiC* and *entC*	CGSC
K. pneumoniae PZH572	Source of *kpdBCD* and *aroY*	ATCC
P. putida KT2440	Source of *catA*	ATCC
Plasmid	**Description**	**Source**
pTrc99A	pBR322 ori, Amp^r, *lacIq*, P_{trc}	Prather Lab, MIT
pTrcCOLAK	ColA ori, Kan^r, *lacIq*, P_{trc}	McKenna et al. (2013)
pY3	p15A, Amp^r, *lacI*, $P_{lac\text{-}UV5}$-*tyrB*-*tyrA*^{fbr}-*aroC* T1-P_{trc}-*aroA*-*aroL*	Juminaga et al. (2012)
pS3	pBBR1 ori; Cm^r, *lacI* $P_{lac\text{-}UV5}$-*aroE*-*aroD*-*aroB*^{op}-*aroG*^{fbr}-*ppsA*-*tktA*	Juminaga et al. (2012)
pKD46	repA101(ts) and R101 ori, Amp^r, *araC, araBp*	CGSC
pCP20	FLP, ts-rep, [cI857](lambda)(ts), Amp^r	CGSC
pTyrAfbr-TutA	*tyrA*^{fbr} of *E. coli* and *tutA* of *C. braakii* inserted to pY3 backbone	This study
pSDC-PchB-EntC	SDC of *T. moniliforme, pchB* of *P. aeruginosa* PAO1, and *entC* of *E. coli* BW25113 inserted to pS3 backbone	This study
pUbiC-Kpd	*ubiC* of *E. coli* and *kpdBCD* of *K. pneumoniae* PZH572 inserted to pTrcCOLAK	Thompson et al. (2015)
pUbiC-PobA	*ubiC* of *E. coli* BW25113 and *pobA* of *P. aeruginosa* PAO1 inserted to pTrc99A	Pugh et al. (2014)
pAroY	*aroY* of *K. pneumoniae* PHZ572 inserted to pTrcCOLAK	Pugh et al. (2014)
pAroY-CatA	*catA* of *P. putida* KT2440 inserted to pQsuB-AroY	This study
pQsuB-AroY	qsuB of C. glutamicum and *aroY* of *K. pneumoniae* PHZ572 inserted to pS3 backbone	This study
pQsuB-AroY-CatA	*catA* of *P. putida* KT2440 inserted to pQsuB-AroY	This study

Table 4.2 Strains Engineered for Phenol and MA Metabolic 'Funneling'

Pathway	Plasmid Combination	Source
PH1	pTyrAfbr-TutA,	This study
PH2	pSDC-PchB-EntC	This study
PH3	pUbiC-Kpd	This study
PHF1	pTyrAfbr-TutA, pSDC-PchB-EntC	This study
PHF2	pTyrAfbr-TutA, pUbiC-Kpd	This study
PHF3	pSDC-PchB-EntC, pUbiC-Kpd	This study
PHF4	pTyrAfbr-TutA, pSDC-PchB-EntC, pUbiC-Kpd	This study
MA1	pQsuB-AroY-CatA	This study
MA5	pUbiC-PobA, pAroY-CatA	This study
MAF	pUbiC-PobA, pQsuB-AroY-CatA	This study

4.2.2. Plasmid Construction

All plasmids used and developed in this study are listed in Table 4.1. All genes were PCR amplified with Q5 High-Fidelity DNA Polymerase (NEB) and a BioRad iCycler, per manufacturer protocols. Custom DNA oligonucleotide primers (See Table C1, Appendix C) were synthesized by Integrated DNA Technologies (IDT, Coralville IA). Genomic DNA (gDNA) templates were prepared using the ZR Fungal/Bacterial DNA MiniPrep kit while plasmid DNA was purified using the Zymo Plasmid MiniPrep kit (both Zymo Research, Irvine CA). Amplified linear DNA fragments were purified using the Zymo DNA Clean & Concentrator MiniPrep kit (Zymo Research). Select purified linear and plasmid DNA were digested using appropriate restriction endonucleases (NEB) and subsequently gel purified using the Zymoclean Gel DNA Recovery MiniPrep kit (Zymo Research). Purified digested DNA fragments were ligated using T4 DNA Ligase (NEB), per manufacturer protocols. Alternatively, purified linear DNA was subsequently used as template DNA for either circular polymerase extension cloning (CPEC) (Quan and Tian 2011) with Q5 High-Fidelity DNA Polymerase according to manufacturer protocols, or Gibson Assembly (Gibson 2009; Gibson et al. 2009) using Gibson Assembly Master Mix (NEB) according to manufacturer protocols. Ligation, CPEC, and Gibson Assembly reactions were transformed into chemically competent *E. coli* NEB10-beta before plating on LB solid agar media supplemented with appropriate antibiotics for selection.

Transformant pools were screened using colony PCR, restriction digest mapping, and finally confirmed by DNA sequencing.

4.2.3. Strain Construction

All strains engineered and used in this study are listed in Table 4.2. Chromosomal in-frame deletions of *pheA* in *E. coli* NST74 was constructed using a modified version of the Datsenko and Wanner method (Datsenko and Wanner 2000), as previously described (Pugh et al. 2014). The *pheA*::FRT-kan^R-FRT, *pykA*::FRT-kan^R-FRT, *pykF*::FRT-kan^R-FRT, and *crr*::FRT-kan^R-FRT deletion cassettes were PCR amplified from *E. coli* JW2580-1, JW1843-2, JW1666-3, and JW2410-1, respectfully. Chromosomal integration of said cassette and subsequent removal of kan^R marker was achieved as previously described (Datsenko and Wanner 2000; Pugh et al. 2014), resulting in the individual construction of *E. coli* NST74 Δ*pheA* and *E. coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF* Δ*crr*.

For phenol, catechol, and MA biosynthesis, *E. coli* NST74 Δ*pheA* and *E. coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF* Δ*crr* were co-transformed with the following combinations of plasmids (note: pathway designation provided in parentheses, see Table 4.2): pTyrAfbr-TutA (PH1), pSDC-PchB-EntC (PH2), pUbiC-Kpd (PH3), pTyrAfbr-TutA and pSDC-PchB-EntC (PHF1), pTyrAfbr-TutA and pUbiC-Kpd (PHF2), pSDC-PchB-EntC and pUbiC-Kpd (PHF3), pTyrAfbr-TutA and pSDC-PchB-EntC and pUbiC-Kpd (PHF4), pQsuB-AroY-CatA (MA1), pUbiC-PobA and pAroY-CatA (MA5), and pUbiC-PobA and pQsuB-AroY-CatA (MAF).

4.2.4. Phenol and MA Production

To investigate the bioproduction of phenol and MA, overnight seed cultures were first prepared in LB and supplemented with 100 mg/L ampicillin, 35 mg/L

kanamycin, and/or 34 mg/L chloramphenicol and used to inoculate (1% vol) 50 mL of MM1 (phenol) or M9M (MA) minimal media supplemented with 20 g/L glucose in 250 mL shake flasks (note: the medium was additionally supplemented with 0.1 g/L Phe when using Δ*pheA*-derived host strains). Shake flask cultures were incubated at 32ºC with shaking at 200 RPM until reaching OD_{600} ~0.7, at which point IPTG induction was performed at a final concentration of 0.4 mM. Cultures were further incubated for a total of up to 120 h, or until significant sugar consumption was no longer detected. Periodically, samples were drawn to measure cell growth (as OD_{600}) as well as sugar and metabolite levels by HPLC analysis, as described below. Prior to centrifugation, samples for MA analysis were first diluted 1:10 with methanol, while samples for Tyr analysis were diluted 1:10 with 1 N HCl and incubated at 55°C for 30 min. All samples were then centrifuged at 11,000 x *g* for 5 min before transferring the supernatant to a glass HPLC vial.

4.2.5. HPLC Metabolite Analysis

Metabolite analysis was performed using a Hewlett Packard 1100 series HPLC system. Separation of Phe, *p*HBA, PCA, salicylate, phenol, catechol, and MA, was achieved using a reverse-phase Hypersil GOLD aQ C18 column (3 mm x 250 mm; Thermo Fisher, Waltham, MA, USA) operated at 45°C with an isocratic 0.8 mL/min mobile phase consisting of 85% (vol.) 5 mM H_2SO_4 and 15% (vol.) acetonitrile. The eluent was monitored using a diode array detector (DAD) set at 215 nm for Phe, salicylate, PCA, and catechol, 260 nm for *p*HBA, and 275 nm for phenol and MA. Separation of Tyr was also achieved on the same Hypersil GOLD aQ C18 column, in this case maintained at 30°C while using a mobile phase consisting of water (A) and methanol plus 0.1% (vol.) formic acid (B) at a constant flow rate of 0.2 mL/min and the following concentration gradient (all by vol.): 5% B from 0 to 8 min, 5% to 40%

B from 8 to 13 min, 40% B from 13 to 16 min, 40 to 5% B from 16 to 21 min, and 5% B from 21 to 31 min. The eluent was monitored using a DAD at 215 nm. Glucose and acetate separation was achieved using an Aminex HPX-87H column (BioRAD, Hercules, CA) operated at 35ºC and detected using a refractive index detector (RID). The column was eluted with 5 mM H_2SO_4 at a constant flow rate of 0.55 mL/min. In all cases, external standards were prepared and used to provide calibrations for concentration determination.

4.3. Results and Discussion

4.3.1. Demonstrating Phenol and MA Production via Synthetic 'Metabolic Funneling'

To avoid the inherent limitations associated with the original phenol and MA biosynthesis routes, a synthetic 'metabolic funneling' approach was investigated as a novel strategy for further improving production metrics. In the case of phenol, multiple 'funneling' strategies were investigated by co-expressing both the original tyrosine-derived pathway (PH1, Figure 4.2) with the salicylate derived pathway (PH2, Figure 4.2) and/or the *p*-hydroxybenzoate (pHBA) derived pathway (PH3, Figure 4.2). In this way, phenol production occurs via either two or three simultaneous and compatible routes (Table 4.3). As with the novel phenol pathways, the 'funneling' derived strains were first introduced and expressed in *E. coli* NST74 Δ*pheA*. In the case of PHF1, titers reached 439 ± 7 mg/L at a glucose yield of 24.0 ± 0.51 mg/g, representing a 16% increase in titer over the best single pathway control, although the yield decreased by 33%. In the case of PHF2, titers reached 355 ± 17 mg/L at a glucose yield of 17.8 ± 1.2 mg/g. Meanwhile, phenol titers by PHF3 reached just 149 ± 10 mg/L at a glucose yield of 8.4 ± 1.0 mg/g. Finally, with all three pathways expressed simultaneously (PHF4), phenol titers reached 205 ± 13 mg/L at a glucose

yield of 11.8 ± 0.72 mg/g. While titers by PHF1 represent a modest increase over the single pathway controls, production metrics achieved by PHF2, PHF3, and PHF4 all represent an overall decrease relative to their single pathway controls. Although the exact mechanism(s) for the observed decrease remain unclear, significant acetate accumulation was observed in all cases (Table 4.3) suggesting high rates of glucose consumption and significant overflow metabolism.

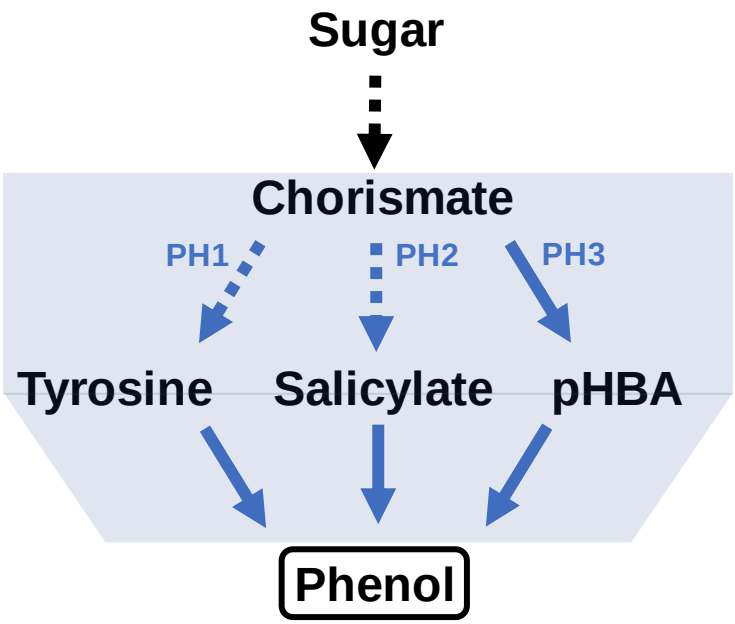


Figure 4.2 Engineering a Synthetic Funnel for Phenol Production. Enhanced biosynthesis can be achieved via the simultaneous co-expression of two or three distinct pathways, all stemming from chorismate as the common endogenous shikimic acid pathway precursor. Black bold dashed arrow represents multiple enzyme reactions native to *E. coli*; blue solid arrows represent individual enzyme steps associated with the pathways engineered in this chapter; pHBA: *p*-hydroxybenzoate.

Meanwhile, to further improve upon the high production potential of the original 3DHS-derived route, a synthetic 'funneling' pathway (MAF) was engineered by co-expressing both the original pathway (MA1, Figure 4.1) and the second, best-performing chorismate-derived pathway (MA5, Figure 4.1), the latter of which provides

an opportunity to 'rescue' additional endogenous precursors not initially 'captured' at the first branch point (i.e., at 3DHS). In addition, this concept importantly avoids the need for multiple nutrient auxotrophies as required for significant production via MA1. As with the phenol and individual MA pathways, MAF was first introduced and expressed in *E. coli* NST74 Δ*pheA*, resulting in the accumulation of 2042 ± 88 mg/L MA at a glucose yield of 102 ± 4.4 mg/g (Table 4.3). When compared to MA1, this suggests that as much as an additional 456 mg/L MA (22% of total MA) was produced via the ability of the 'lower branch' (i.e., MA5) to assimilate precursors (in this case, chorismate) not initially taken in via the 'upper branch' (i.e., MA1). Overall, when compared to strains expressing the single, parent pathways (i.e., MA1 or MA5 alone), the combined 'funneling' strategy of MAF resulted in MA titer improvements of 29 and 150%, respectively, and 29 and 132% in terms of achievable yields. However, similar to the phenol production, significant acetate accumulation was observed (Table 4.3), again suggesting high rates of glucose uptake and overflow metabolism.

Table 4.3 Evaluating Metabolic 'Funneling' for Phenol and MA Biosynthesis. The use of synthetic 'metabolic funneling' for improving production was investigated using *E. coli* NST74 Δ*pheA* as the host background. Strains were cultured for 120h. Error represents one standard deviation from triplicate experiments. MA: Muconic acid.

Pathway	Product (mg/L)	Glucose Utilization (%)	Acetate (g/L)	$Y_{P/S}$ (mg/g)
PH1	Phenol: 377 ± 23	100 ± 1	9 ± 0.1	18.7 ± 0.72
PH2	Phenol: 377 ± 14	53 ± 0.7	6 ± 0.3	35.7 ± 0.80
PH3	Phenol: 149 ± 12	100 ± 1	4 ± 0.3	7.3 ± 0.31
PHF1	Phenol: 439 ± 7	90 ± 0.9	8 ± 0.2	24.0 ± 0.51
PHF2	Phenol: 355 ± 17	100 ± 1	6 ± 0.2	17.8 ± 1.2
PHF3	Phenol: 149 ± 10	90 ± 0.4	6 ± 0.3	8.4 ± 1.0
PHF4	Phenol: 205 ± 13	87 ± 0.3	5 ± 0.2	11.8 ± 0.72
MA1	MA: 1586 ± 11	100 ± 1	12 ± 0.2	79.3 ± 0.53
MA5	MA: 819 ± 44	100 ± 1	12 ± 0.2	40.9 ± 2.2
MAF	MA: 2042 ± 88	100 ± 1	11 ± 0.2	102 ± 4.4

4.3.2. Host Engineering to Enhance Precursor Availability

To further improve phenol and MA production, subsequent strain engineering efforts were focused on increasing total carbon flux into the shikimic acid pathway and reducing overflow metabolism observed with high glucose uptake rates (Liu et al. 2014). Flux into the shikimic acid pathway is initially controlled by 3-deoxy D-arabinoheptulose 7-phosphate (DAHP) synthase, whose two substrates are phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) (Bongaerts et al. 2001; Gosset 2009; Rodriguez et al. 2014). As previously demonstrated (Gosset 2005; Postma et al. 1993), increasing the intracellular availability of PEP is an effective strategy for enhancing the production of aromatic amino acids (Liu et al. 2014) and other products from intermediates of the shikimic acid pathway (Noda et al. 2016). In glucose-fed cultures, PEP availability can be increased by blocking its conversion to pyruvate via deletion of *pykA* and *pykF*, both of which encode isozymes of pyruvate kinase. Meanwhile, rapid uptake of glucose has been previously reported to result in the accumulation of acetate (Gosset 2005) – an undesirable byproduct which can ultimately inhibit cell metabolism (Shiloach et al. 1996) and result in lower aromatic product yields (Liu et al. 2014). As seen in Table 4.3, significant levels of residual acetate were observed with phenol and MA production, with levels reaching as high as 12 g/L. Carbohydrate repression resistant null mutants (i.e., Δ*crr*) have been previously shown to display lower rates of glucose uptake and thus reduced overflow metabolism – a strategy demonstrated as effective for enhancing phenylalanine production (Liu et al. 2014). Accordingly, *E coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF* Δ*crr* was next constructed and evaluated as a phenol and MA production host.

Table 4.4 Comparing Engineering Strategies for Phenol and MA Production. The use of synthetic 'metabolic funneling' for improving production was investigated using *E. coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF* Δ*crr* as the host background. Strains were cultured for 120h and consumed 20 g/L glucose. Error represents one standard deviation from triplicate experiments.

Pathway	Phenol (mg/L)	Glucose Utilization (%)	Acetate (g/L)	$Y_{P/S}$ (mg/g)
PH1	329 ± 9	100 ± 1	6 ± 0.1	16.4 ± 0.65
PH2	277 ± 15	100 ± 1	7 ± 0.1	13.8 ± 0.19
PHF1	575 ± 19	100 ± 1	5 ± 0.3	28.8 ± 0.34
MA1	1792 ± 28	100 ± 1	n.d.	89.6 ± 1.4
MAF	3153 ± 149	100 ± 1	7 ± 0.1	158 ± 7.4

In the case of phenol, the pathway focus narrowed to the best performing pathways: PH1, PH3, and PHF1. The individual pathways PH1 and PH3 displayed a modest decrease in titer and yield when using *E coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF Δcrr* as host, with final phenol titers reaching 329 ± 9 and 277 ± 15 mg/L, respectively. On the other hand, phenol production by PHF1 was slightly enhanced, with final titers reaching 575 ± 19 at a glucose yield of 28.8 ± 0.34 – a 1.3 and 1.2-fold increase relative to the previous generation strain. In addition, the titer demonstrated by the novel 'metabolic funneling' pathway represents a 4-fold increase over the original phenol biosynthesis reported using the solvent tolerant *P. putida* (Wierckx et al. 2005) and a 1.4-fold increase over the highest *E. coli* derived phenol production reported to date (Kim et al. 2014). In both cases, acetate accumulation was reduced relative to the same pathways used in *E coli* NST74 Δ*pheA*, although significant levels remained at the end of culturing (as much as 6 g/L, Table 4.4). Both engineered pathways each result in the generation of 1 additional pyruvate via TPL (PH1, step J, Figure 3.1) and chorismate pyruvate lyase (encoded by *pchB*, PH2, step H, Figure 3.1) activities. While future works will be required to test this hypothesis, it is possible that said pyruvate is then converted to acetyl-CoA and ultimately to acetate via *E. coli*'s phosphate acetyltransferase and acetate/acetyl-CoA pathway (encoded by *pta-ackA*), providing

the cell with 1 ATP in the process. Further host engineering to incorporate additional mutations previously reported as effective for reducing acetate accumulation during aromatic amino acid biosynthesis will also be investigated (Castaño-Cerezo et al. 2009; Liu et al. 2016; Wang et al. 2013; Wolfe 2005) with the goal of further improving phenol production.

Meanwhile, in the case of MA, the pathway focus narrowed to just MA1 and MAF: the two most promising pathways. As seen in Table 4.4, acetate accumulation was effectively eliminated in the case of MA1, while the additional mutations enabled a modest (~13%) increase in MA production with final titers reaching 1792 ± 28 mg/L at a glucose yield of 89.6 ± 1.4 mg/g. On the other hand, MA production by MAF was significantly enhanced using *E coli* NST74 Δ*pheA* Δ*pykA* Δ*pykF Δcrr* as host, with final MA titers reaching 3153 ± 149 mg/L at a glucose yield of 158 ± 7.4 mg/g – both increase of over 1.5-fold increase relative to the use of *E coli* NST74 Δ*pheA* as host, and the highest production metrics achieved in this study. Furthermore, this maximal titer also is 31% higher than the 2.4 g/L of MA reported by Draths and Frost via MA1 in an *E. coli ΔaroE* background (Draths and Frost 1994), and was notably achieved while generating only a single auxotrophy (i.e., for Phe). Further comparing the results of Tables 4.3 and 4.4, although the apparent flux of precursor through the 'upper branch' (i.e., MA1) was only slightly improved, the 'lower branch' (i.e., MA5) appeared to offer an even greater benefit, in this case enabling the production of as much as an additional 1361 mg/L MA (43% of total MA produced). The additional Δ*pykA*, Δ*pykF*, and *Δcrr* mutations appear to have more greatly improved the intracellular availability of chorismate than that of 3DHS. Interestingly, acetate accumulation reemerged with MAF, reaching up to 7 ± 0.1 g/L by the end of the culture (Table 4.4). Unlike MA1, the first step of MA5 (step D, Figure 3.1) is catalyzed by chorismate pyruvate lyase (encoded by native *ubiC*) and results in the generation of 1 pyruvate molecule, which,

similar to the phenol results, may ultimately be converted to acetate. The same host engineering and additional mutations previously discussed could similarly be investigated with the goal of reducing acetate accumulation and improving MA production via MAF.

4.3.3. Synthetic 'Funneling' for Renewable Chemical Production

Two novel 'metabolic funneling' pathways were constructed for the complete biosynthesis of phenol and MA from glucose, resulting in titers reaching 0.58 and 3.1 g/L, respectively. In both cases, additional strain engineering to reduce unwanted acetate accumulation may ultimately lead to higher achievable production metrics. Whereas product toxicity is a common concern with the production of aromatic chemicals (McKenna and Nielsen 2011; McKenna et al. 2013), further increases in phenol production may ultimately reach a toxic threshold (estimated at 1.8 g/L in rich media (Kim et al. 2014)). Others have demonstrated that *in situ* product removal can effectively improve overall phenol biosynthesis (Miao et al. 2015b); such an approach may prove analogously helpful with the phenol 'funneling' pathway combination.

Although many products of interest have only one heterologous pathway reported to date, one interesting prospect for future works is the biosynthesis of 3-hydroxypropionate (3HP). 3HP is a versatile platform chemical used for the production of polymers, acrylates, and lubricants (Kumar et al. 2013). To date, no fewer than seven novel routes have been proposed, all of which ultimately stem from either pyruvate or PEP (Table 4.5). Of the seven proposed pathways, only three (Pathways 2, 3, and 6, Table 4.5) are reported to be energetically favorable, with either a 0 or +1 net ATP yield per mol of 3HP produced (Kumar et al. 2013). Given these three favorable pathways, all combinations could be investigated for enhanced 3HP production using a synthetic 'funneling' strategy. Furthermore, an interesting

alternative approach would be to investigate the simultaneous co-expression of at least one energetically favorable pathway in parallel with at least one unfavorable pathway. More specifically, Pathway 6 is energetically favorable with a net ATP yield per mol 3HP of +1 and has been demonstrated to result in the highest achievable titer and yield reported to date (Henry et al. 2010; Jiang et al. 2009). Meanwhile, Pathway 4 has a net ATP yield per mol 3HP of -1 and has been demonstrated as functional, although 3HP was produced at low levels (Liao et al. 2007). Future works could investigate metabolic 'funneling' as a strategy for balancing pathway energetics, in the case utilizing the extra ATP produced via Pathway 6 to supplement the deficiency created via Pathway 4.

Table4.5 3-hydroxypropionate Biosynthesis Pathways. Net ATP is dependent on endogenous substrate and enzyme selection. [a]Kumar et al. 2013

Pathway	Intermediates	Net ATP (mol / mol 3HP)[a]
1	Pyruvate → Lacate → Lactoyl-CoA → Acryloyl-CoA → 3HP-CoA → 3HP	1/0
2	Pyruvate → Acetyl-CoA → Malonyl-CoA → 3-oxopropanoate → 3HP	0
3	Pyruvate/PEP → OAA → Aspartate → β-alanine → 3-oxopropanoate → 3HP	1/0
4	Pyruvate/PEP → OAA → Aspartate → β-alanine → β-alanyl-CoA → Acryloyl-CoA → 3HP-CoA → 3HP	0/-1
5	Pyruvate/PEP → Succinate → Propionate/Propionyl-CoA → Acryloyl-CoA → 3HP-CoA → 3HP	-1/-0.33
6	Pyruvate → α-alanine → β-alanine → 3-oxopropanoate → 3HP	1
7	Pyruvate → α-alanine → β-alanine → β-alanyl-CoA → Acryloyl-CoA → 3HP-CoA → 3HP	1/0

4.4. Conclusions

A synthetic 'funneling' approach was investigated to increase phenol and MA production while avoiding the inherent limitations associated with each original biosynthesis route. In the case of phenol, co-expression of two phenol pathways in parallel enabled titers exceeding 0.5 g/L in shake flask cultures. Meanwhile, co-expression of two MA pathways in parallel resulted in a novel 'funneling' pathway enabled MA production at titers >3 g/L in shake flask cultures. Utilizing said approach is expected to have broad applications metabolic engineering for improving the biosynthesis of an array of products of interest.

5. DISCUSSION AND FUTURE WORK

Abstract

Novel phenol and muconic acid biosynthesis pathways have successfully been engineered to overcome inherent limitations associated with the original biosynthesis routes. Despite these novel contributions, significant room for improvement remains to further increase production metrics. This chapter discusses potential future work to further increase phenol and muconic acid production by characterizing and utilizing various dynamic flux control strategies with the potential to greatly increase availability of the relevant pathway precursors.

5.1. Engineering a Dynamic Phenotype Switch

Through the rational design and engineering of natural and non-natural metabolic pathways, *E. coli* can serve as an efficient biocatalyst for the transformation of renewable sugars to various bulk and fine chemicals (Carothers et al. 2009; Curran and Alper 2012; Lee et al. 2011; Martin et al. 2009; Xie et al. 2006). Although developments in metabolic engineering and synthetic biology continue to expand an ever-growing library of heterologous pathway 'parts', there remains significant room for improvement regarding overall pathway performance. More specifically, in addition to producing the targeted end-product of interest, undesirable competing metabolic flux toward biomass and essential nutrients production remains a significant drain for available carbon resources (Figure 5.1) (Mao et al. 2009; Weikert et al. 1998; Youngquist et al. 2013). Despite the identification of various native flux control mechanisms, including conditional promoter induction, riboswitches, attenuation mechanisms, and feedback regulation, these strategies are difficult to predict and therefore remain challenging to engineer.

Whereas maximum product yields are achieved when competing metabolic flux is minimized, the engineering of static and dynamic control mechanisms to efficiently regulate both engineered and competing pathway flux is essential for increasing production metrics (Mao et al. 2009; Weikert et al. 1998; Youngquist et al. 2013). Static mechanisms include modulating gene expression levels through the rational adjustment of RBS strength, promoter strength, promoter induction level, and plasmid copy number. While these techniques offer fine-tunable control of gene expression, they offer minimal impact with respect to restricting undesirable carbon loss to competing pathways. In contrast, gene deletions eliminate carbon loss to competing metabolic networks, although this strategy is limited to either non- or only conditionally-essential genes. Whereas the former may have no discernable

phenotypic impact, the latter typically requires additional media supplements to support cell viability resulting in reduced host fitness and poor scalability.

In contrast to the above mentioned, static strategies, competing pathway flux and undesired carbon loss can be minimized by engineering dynamic control strategies, resulting in more tractable and higher efficiency production platforms. For example, sRNA has been used in pathway engineering to target non-essential genes as part of enhanced tyrosine and cadaverine biosynthesis (Na et al. 2013). Meanwhile, a metabolic toggle switch variant was previously demonstrated to effectively modulate flux between growth essential pathways and the production of 3-hydroxypropionic acid (Tsuruno et al. 2015). Moreover, by constructing a controllable native protein degradation system, myo-inositol production was enhanced by modulating the key glycolytic enzyme Pfk-I (Brockman and Prather 2015). Finally, although dynamic chromosomal deletions have been modeled in silico, this approach has yet to be

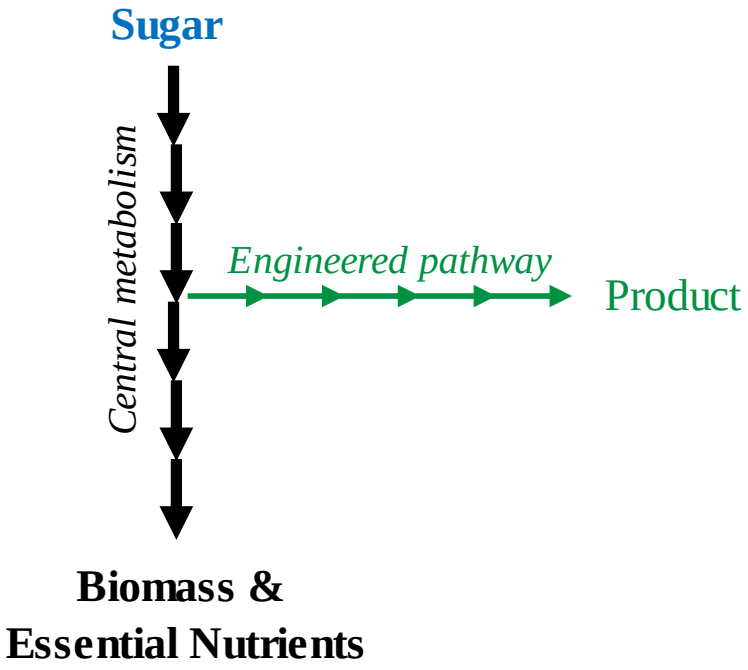


Figure 5.1 Metabolic Flux Distribution for Enhanced Product Biosynthesis

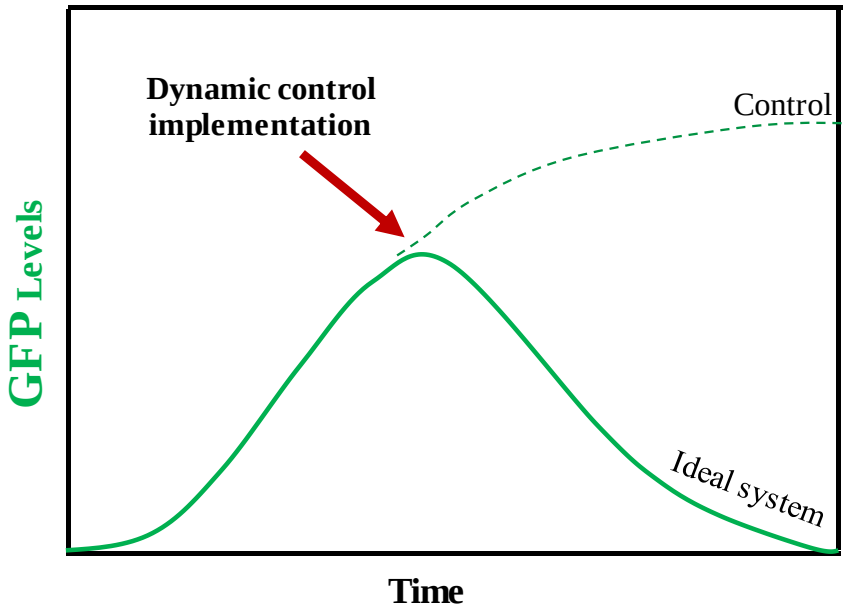


Figure 5.2 Dynamic Flux Control Response. Flux control mechanisms can be evaluated using GFP as a model system. Various GFP expression repression techniques can be investigated. The ideal outcome (solid line) will result in the rapid repression of GFP expression relative to the control (dashed line).

demonstrated for pathway engineering (Gu et al. 2012). Despite the successful implementation of these techniques, a head-to-head evaluation clearly identifying the relative merits and pitfalls associated with each strategy has not been performed. Furthermore, said strategies have typically been applied for the enhanced biosynthesis of one product with no further application to disparate pathways and products, resulting in underutilized techniques and a general ambiguity regarding their efficacy.

The overall goal of future work to research presented here is to engineer, characterize, and comparatively evaluate five unique mechanisms, and combinations thereof, for creating a dynamic phenotype switch. These mechanisms will target transcriptional, translational, and post-translational control to achieve the desired dynamic phenotype switch. GFP will serve as an initial model while the relative and absolute performance of each design can be compared. In all cases, the ideal control system will result in rapid loss of GFP expression (i.e., minutes vs. hours) and

subsequent depletion of detectable GFP levels (Figure 5.2). Each approach proposed for a dynamic phenotype switch has its own inherent advantages and disadvantages (Table 5.1), as will be discussed in greater detail in the following sections. Overall, the most effective control system design may require simultaneous deployment of multiple mechanisms considered (e.g., CRISPRi in conjunction with targeted protein degradation) to ultimately realize both rapid and efficient phenotypic switching. The individual mechanisms of interest, as well as relevant factors to be considered with respect to their implementation and design, are discussed in the subsequent sections.

Table 5.1 Dynamic Phenotype Control Mechanisms. Engineered approaches to be investigated each have unique advantanges and disadvantages.

Dynamic Approach	Advantages	Disadvantages
CRISPRi	• Targets DNA (single copy) • Multiple gene targets simultaneously	• Two-part system • Off-target effects • Requires expression of heterologous deactivated protease • Residual protein
Metabolic toggle switch	• Tunable control between growth and product formation • Reversible	• Large construct to implement • Inherently leaky expression • Reversible
Dynamic chromosomal deletion	• Irreversible	• Requires chromosomal replacement • Difficult to fine tune • Irreversible
sRNA	• No host engineering • Small construct (~ 300 bp per target) • Multiple gene targets simultaneously • No additional protein expression required	• Targets mRNA (multiple copies) • Off-target effects • Requires fine tuning for adequate repression • Residual protein
Protein degradation	• Fast degradation • No residual protein	• Specific protein tag required • Require expression of exogenous or tuning of endogenous protease • Not all proteins degradable

5.1.1. Transcriptional Control

To achieve the ideal control system, the use of transcriptional control will first be investigated via the clustered regularly interspaced short palindromic repeats (CRISPR) system: a system that has emerged as a powerful tool for a variety of applications, including gene editing and gene silencing (Jiang et al. 2015). In nature, these systems typically operate as a defense mechanism to cleave non-native DNA. CRISPR systems are comprised of two essential components: (a) a small guide RNA (sgRNA) and (b) the Cas9 protein. The sgRNA consists of a base pairing region, a Cas9 handle, and a terminator, and is responsible for pairing to a specific DNA locus complimentary to the base pairing region. Once bound to the DNA locus, the Cas9 protein cleaves the double strand DNA. However, a deactivated Cas9 protein (dCas9) protein is catalytically inactive and does not exhibit endonuclease activity (Hawkins et al. 2015). Once bound to the DNA locus, the sgRNA-dCas9 complex precludes RNA polymerase binding, thereby disrupting gene transcription. Notable advantages of this technique include the transcriptional repression of multiple gene targets simultaneously, whereas disadvantages include off-target effects and the requirement of an additional heterologous protease (Table 5.1). To investigate this approach, a synthetic CRISPRi system will be constructed capable of modulating GFP transcription levels (Figure 5.3). As an initial test platform, a custom designed CRISPRi system (Figure 5.3A), composed of a custom synthesized sgRNA and dCas9, will be used for transcriptional repression of GFP (Figure 5.3B). GFP expression levels over time will be evaluated to fully characterize the inducible CRISPRi system.

The second transcriptional control approach will investigate the use of a metabolic toggle switch. One prospective toggle switch design relies on a dual promoter-repressor system that can readily alternate between transcription of two genes (or operons), dependent on the inducer present in the media (Litcofsky et al.

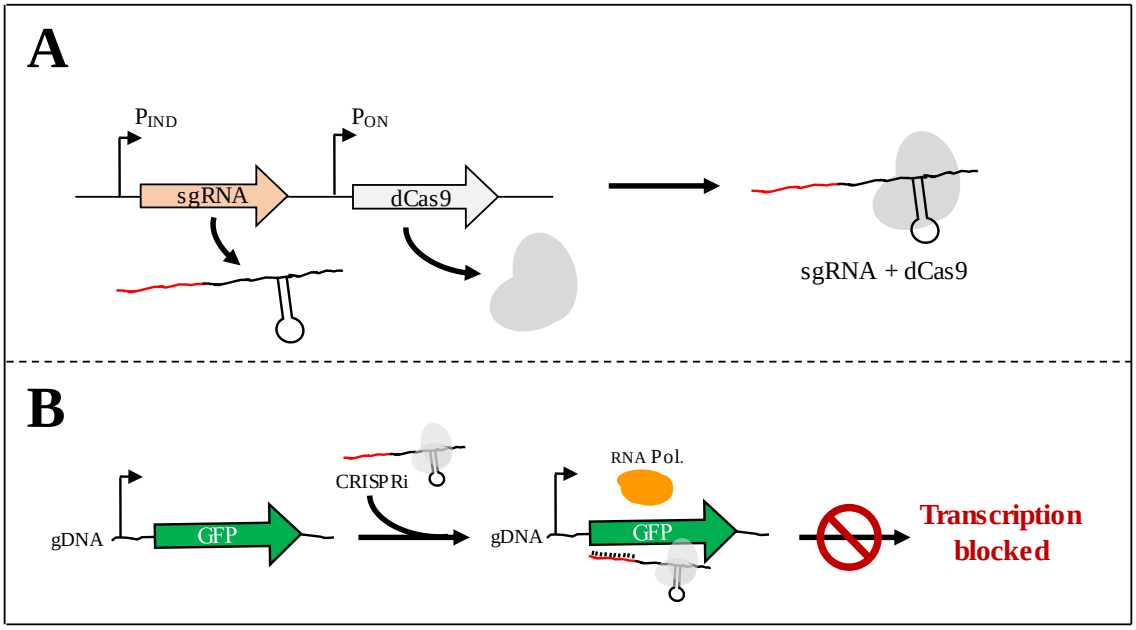


Figure 5.3 Synthetic CRISPRi**.** (A) Custom designed circuits will be engineered for controllable sgRNA and dCas9 expression. (B) Synthetic CRISPRi circuits will be tested for GFP transcriptional repression.

2012). Advantages of this approach include the ability to fine tune expression levels and general reversibility, although such constructs are often large and inherently leaky. To investigate this approach, a dual GFP-RFP construct will be evaluated (Figure 5.4A). Said system relies on the use of a previously developed toggle switch that can readily alternate between GFP and RFP expression. In the absence of inducer (aTc), RFP transcription will be 'on' whereas GFP transcription will be 'off'. Following addition of inducer (aTc), RFP transcription will be 'off' and GFP transcription will be 'on' (Figure 5.4B). RFP and GFP expression over time will be evaluated to fully characterize the toggle switch system.

The last transcriptional control approach to be investigated is the implementation of a dynamic chromosomal deletion cassette, modeled after the traditional chromosomal gene deletion strategy (Figure 5.5) (Datsenko and Wanner 2000). Traditional chromosomal gene deletions (Figure 5.5A) typically result in the greatest overall impact to increasing titer and yield by eliminating competing pathway

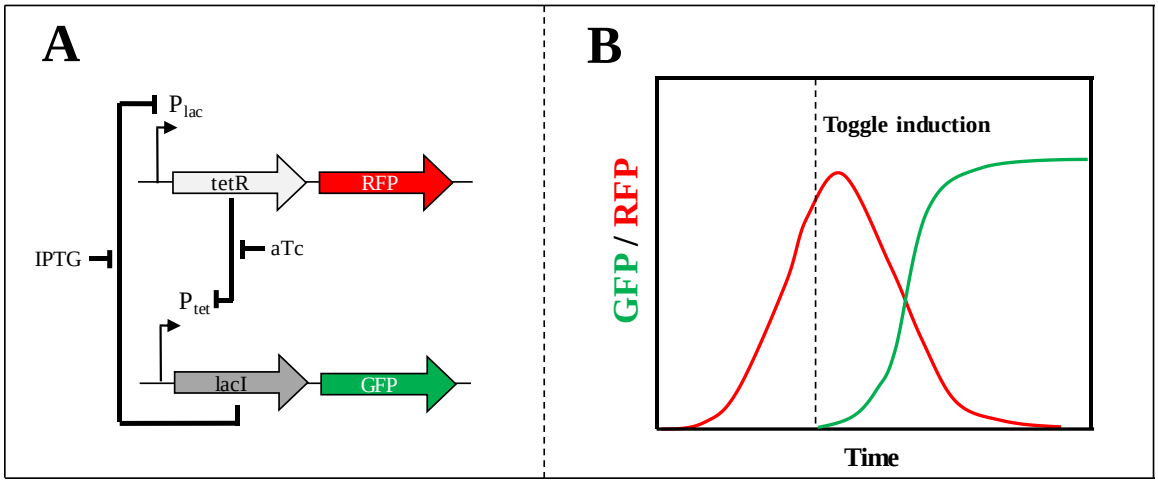


Figure 5.4 Engineered Toggle Switch. (A) Metabolic toggle switch used to impart a sharp and efficient phenotype switch. (B) Ideal phenotype switch output. In the absence of aTc, cells express RFP. Following addition of aTc, a phenotype switch is observed by the loss of RFP expression and presence of GFP expression.

flux. Gene deletions are most commonly constructed via homologous recombination using a selectable marker (kanamycin or chloramphenicol resistance) flanked by FRT sites. Although this method is well established and highly effective at reducing competing pathway flux, gene deletions remain limited to non- and conditionally-essential targets. Expanding on this existing technology, the engineering and implementation of a dynamic chromosomal deletion system is proposed to impart flux control (Figure 5.5B). Gene targets of interest, in this case GFP, will be chromosomally flanked by FRT sites. Following expression of FLP recombinase, the FRT flanked gene target will be chromosomally deleted, resulting in a controllable phenotype switch. Whereas the original gene target must effectively be replaced with the FRT flanked variant, this strategy is inherently irreversible and therefore offers additional engineered microorganism biocontainment as the loss of an essential gene renders the host unviable at the end of culturing. Analogous to the previously described approaches, GFP levels over time will be evaluated to characterize this phenotype switch.

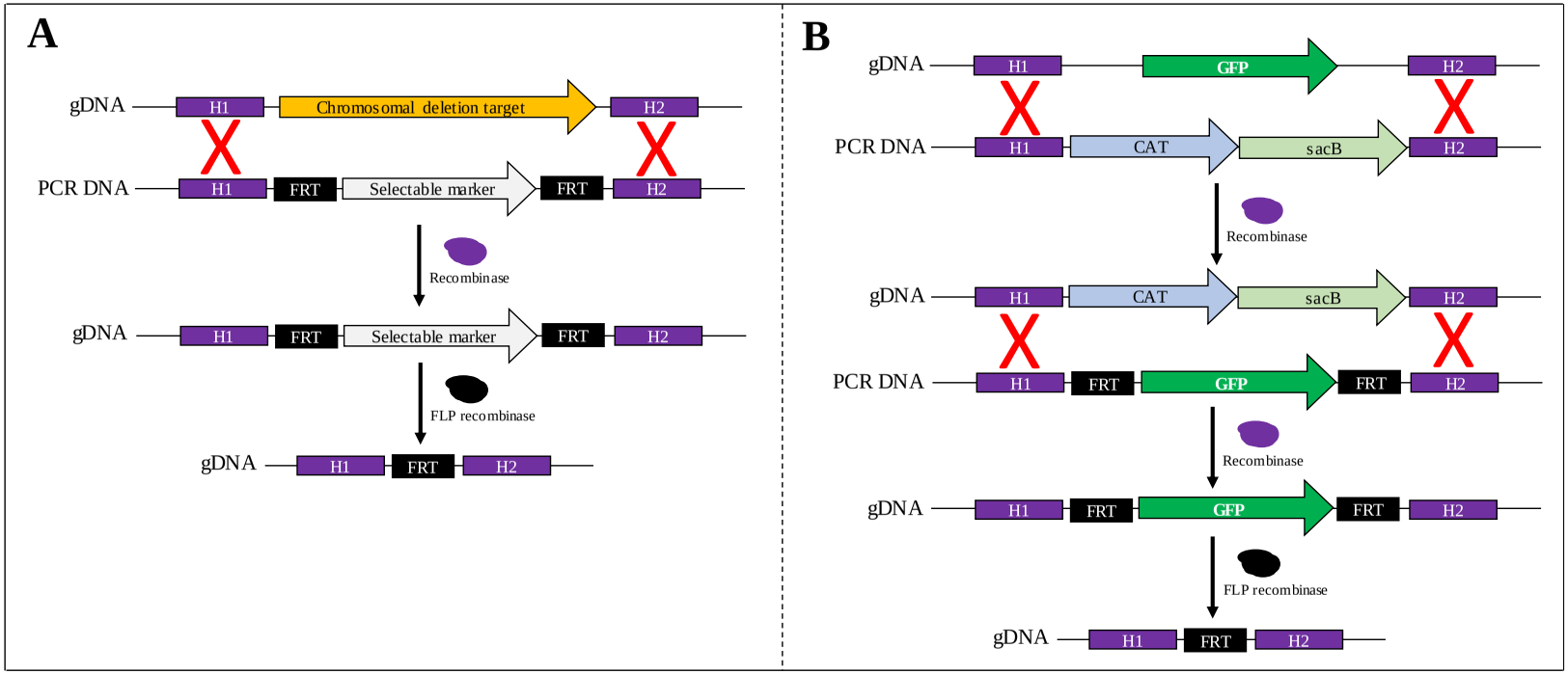


Figure 5.5 Engineering Dynamic Homologous Recombination. (A) Traditional chromosomal deletion using homologous recombination. A linear amplified deletion cassette is constructed to contain chromosomal homology upstream and downstream of the target gene. A selectable marker is integrated using in place of the gene target using an inducible recombinase. Marker excision is achieved via an inducible FLP recombinase, resulting in a complete chromosomal deletion. (B) Engineering a dynamic chromosomal deletion. Chromosomal GFP will be replaced with a FRT flanked variant using a scarless homologous recombination approach. Following dynamic FLP recombinase induction, GFP will be excised from the chromosome.

5.1.2. Translational Control

E. coli relies on a variety of RNA-mediated mechanisms to modulate gene expression through translational control, including attenuation mechanisms, riboswitches, and *cis*- and *trans*-encoded sRNA. Both attenuation mechanisms and riboswitches rely on complex secondary RNA self-interactions and/or metabolite binding under select environmental conditions. Whereas both methods effectively preclude ribosomal attachment and subsequent translation, these control mechanisms are difficult to design for metabolic engineering applications. Meanwhile, *cis*-encoded sRNA (csRNA) is a small regulatory RNA transcribed from the sense DNA strand near the target gene. Because of the proximity, csRNA typically consists of lengthy target binding regions (~75 nucleotides) and secondary structures (up to 500 bp), resulting in complete sequences that overlap with the coding region of interest. Although effective at blocking ribosomal binding, csRNA is therefore difficult to design and construct for metabolic engineering applications due to the chromosomal location. Finally, *trans*-encoded sRNA is transcribed at a distant location relative to the target gene. As these sRNA molecules are transcribed at various locations throughout the chromosome, each trans-encoded sRNA can have multiple gene targets, dependent on the target binding region. Accordingly, *trans*¬-encoded sRNA target binding sequences are readily predictable and are an ideal tool for modulating gene expression through translational control. Although the sRNA constructs are small and require no additional host engineering, calculated fine tuning is required with added caution towards potential off-target effects. Accordingly, custom sRNA circuits will be designed to modulate GFP translational levels (Figure 5.6). sRNA expression will be under control of an inducible promoter (Figure 5.6A). One present, sRNA scaffolds are stabilized by native Hfq protein. Stable sRNA scaffolds will bind to the mRNA at a pre-determined location, resulting in translational inhibition (Figure 5.6B). GFP levels will be evaluated

over-time to characterize the response and efficacy of sRNA as a phenotype control switch.

5.1.3. Post-translational Control

To demonstrate and characterize the use of post-translational control as a phenotype switch, tunable protein degradation will lastly be investigated. *E. coli* contains several endogenous proteases that degrade select proteins labeled with specific amino acid tails. Whereas variants of the amino acid tails result in dissimilar degradation rates, inducible control cannot be achieved without extensive host engineering. More recently, a tunable protein degradation system that does not rely on native protease activity was developed in *E. coli (Cameron and Collins 2014)*. The system is based on a *Mesoplasma florum* transfer-messenger RNA system that allows for steady-state protein levels with tunable degradation via Lon protease (*mf*-Lon),

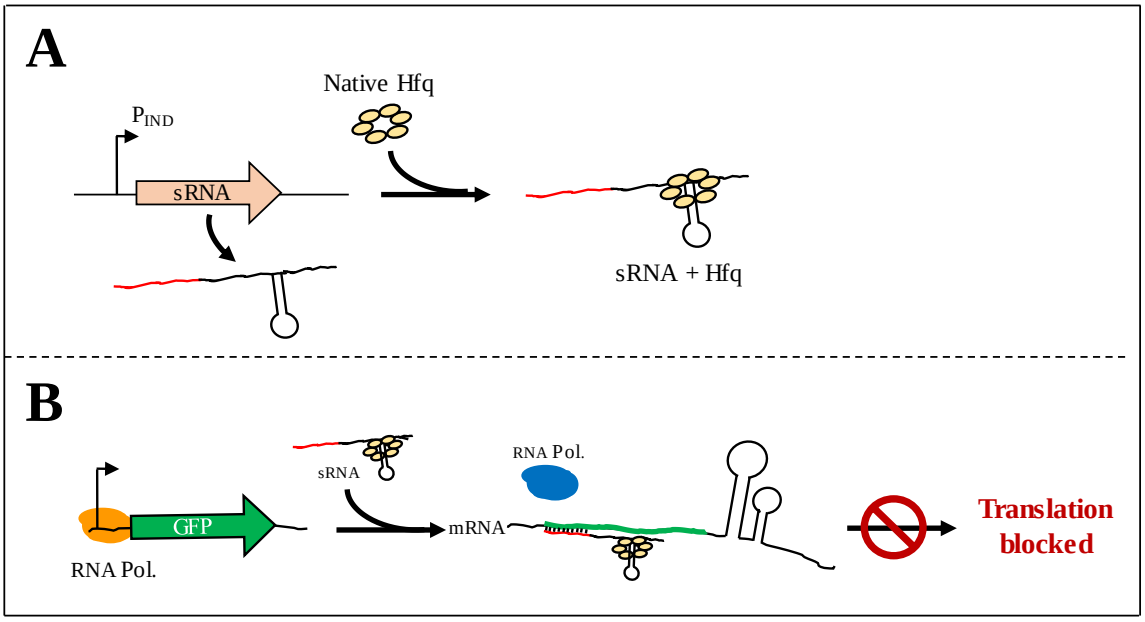


Figure 5.6 Small Synthetic Regulatory RNA (sRNA). (A) Synthetic sRNA circuits will be engineered for custom designed sRNA expression. sRNA transcripts are stabilized by native Hfq protein binding. (B) Synthetic sRNA constructs will be evaluated for the translational repression of GFP.

does not require disruption of the host organism's native protein degradation mechanisms, and can be seamlessly integrated for the targeted proteolysis of protein variants containing a specific degradation tag. To demonstrate this approach as a dynamic phenotype switch, native GFP will be replaced with a custom tagged GFP variant and evaluated for degradation following expression of the *mf*-Lon protease (Figure 5.7). Successful proteolysis will be determined by evaluating GFP expression levels following *mf*-Lon protease induction.

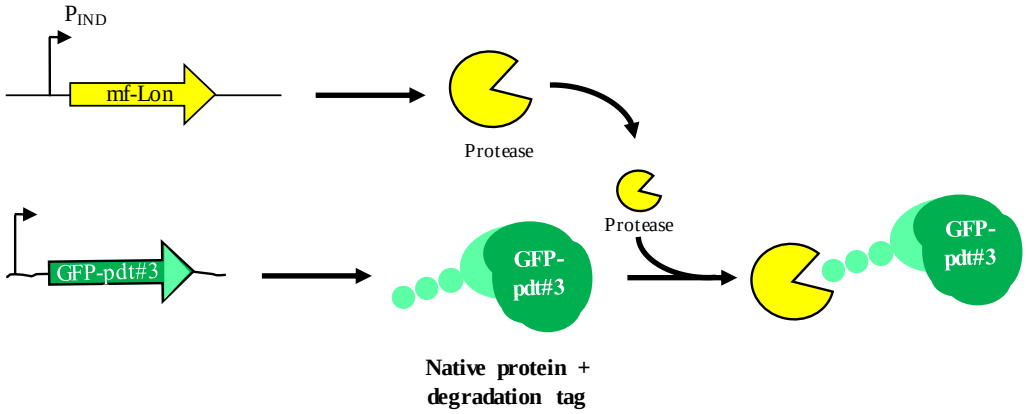


Figure 5.7 Tunable Protein Degradation. *mf*-Lon mediated protein degradation achieved via custom tagged proteins containing a custom designed C-terminus amino acid tail (e.g. pdt#3).

5.1.4. Phenotype Switch Design Considerations

To achieve complete repression of GFP expression, several design considerations will be implemented. In the case of CRISPRi, an important design element is the sequence of the custom sgRNA. Here, a 20-25 nucleotide base-pairing region will be used as this has been found to be sufficient for high binding energy with the target nucleotide sequence. For example, when co-expressed with dCas9, an sgRNA with 20 nucleotides of complementarity has been shown to result in up to 99.9% transcriptional repression. Repression of transcription has also been reported

to be the strongest when the selected binding site is positioned close to or on the transcription start site of the target gene. In this case, several binding locations will be comparatively evaluated to identify that which confers the greatest repression efficiency. Finally, the base-pairing region should also be designed to minimize off-target binding as such occurrences can unpredictably impact host function. *In silico* design tools (e.g. E-CRISP, sgRNAcas9) will be used to help ensure target site specificity. With these factors in mind, a comprehensive pool of candidate sgRNA sequences will be designed, synthesized, and comparatively screened (further described below) to identify the most effective design strategy. Preliminary data using CRISPRi for MA production is contained in Appendix D.

In the case of the metabolic toggle switch, an important design element is the complimentary promoter-repressor selection and induction level. Here, a previously developed toggle switch will be used to effectively alternate between RFP and GFP expression. Our system is comprised of a P_{lac}-tetR-RFP and P_{tet}-lacI-GFP promoter-repressor system. In the absence of aTc, this system expresses tetR-RFP, thereby repressing lacI-GFP expression. Following addition of aTc, tetR-RFP expression is repressed, allowing for lacI-GFP expression. Furthermore, addition of IPTG can return the toggle switch to the original state (i.e. tetR-RFP expression, lacI-GFP repression). Accordingly, varying inducer levels (both aTc and IPTG) will be comparatively evaluated to identify that which confers complete expression and switch between RFP and GFP. Preliminary data using a metabolic toggle switch for MA production is contained in Appendix D.

Next, an important design element for the dynamic chromosomal deletion approach is expression of the FLP recombinase. In this case, *in silico* models have demonstrated that the FLP recombinase can reach maximum DNA excision levels approaching 100%. To achieve this high level of excision, promoter selection and

corresponding inducer levels will be investigated. In addition, plasmid copy number will be varied and comparatively evaluated to achieve as close to 100% excision as possible.

In the case of sRNA, a custom designed nucleotide base-pairing region must confer high binding energy with the target nucleotide sequence to result in adequate translational repression. The use of sRNA to dynamically control pathway flux has previously been implemented for the enhanced biosynthesis of tyrosine and cadaverine, resulting in the production of 2 g/L tyrosine and a 55% increase in cadaverine production relative to the control. In this case, a 20-30 nucleotide region and a target binding energy between -30 and -40 kcal/mol was required to achieve greater than 95% translational repression of select gene targets. Levels of translational repression were further dependent on the location of the binding site relative to the ribosomal binding site. In this case, several binding energies and locations will be comparatively evaluated (as determined using the *in silico* design tool DINAmelt) to identify the combination that results in the highest translational repression. Finally, the target binding region should also be designed to reduce unwanted off-target effects. Here, off target binding energies less than -10 kcal/mol have been reported to have no ill effect on overall host function. The *in silico* design tool TargetRNA2 will be used to ensure target binding specificity. With these factors in mind, a comprehensive pool of sRNA sequences will be comparatively evaluated to determine the most efficient system for translational control.

Finally, tunable post-translational control will be investigated by evaluating GFP degradation rates. In this case, a tagged GFP variant that results in rapid and efficient degradation in the presence of exogenous *mf*-Lon protease will be constructed. Tagged variants have previously been reported to be stable in the absence of *mf*-Lon protease, while levels of degradation are controlled by the 27 amino acid tail added to

the C-terminus of the protein. In this case, the previously reported pdt-#3 tag will be investigated first, as this sequence resulted the in rapid (< 6 hours) and complete loss of the targeted protein. Additional rates and levels of protein degradation will be investigated by altering the C-terminus tag sequence. In addition, various promoter strengths and plasmid copy number driving *mf*-Lon protease expression will be investigated to fine tune degradation rates and levels.

5.1.5. Phenotype Switch Characterization

As mentioned previously, the ideal phenotype control system will result in the rapid loss of GFP expression and subsequent depletion of detectable GFP levels. Accordingly, the *in vivo* function of all proposed dynamic phenotype switches will first be comprehensively examined with respect to their kinetic and equilibrium behaviors using a suite of analytical tools. Using the initial GFP-based systems, a dynamic phenotype switch capable of a rapid and sharp phenotype transition will be identified. In addition to the design considerations outlined above, several other relevant factors will also be explored to first understand and then optimize their effect. In addition to culture temperature, the role of induction strength and timing (i.e. exponential vs. stationary phase) will be of interest. A preliminary assessment of the function and efficacy of the different dynamic phenotype switches will easily be gained from monitoring the relative GFP expression levels via a custom designed plate reader assay and/or flow cytometry. Efficient repression of GFP expression (i.e. transcriptional and translational repression) will result in steady-state GFP levels equivalent to those observed at the time of dynamic phenotype switch induction. Efficient loss of GFP expression (i.e. post-translational repression) will result in the complete loss of detectable GFP levels. More detailed analyses will subsequently be performed for the most promising phenotype switch designs (i.e. sgRNA sequences, sRNA sequences,

degradation tag, etc.) and conditions (i.e. induction strength and timing). Quantitative real-time PCR will be used to determine the rate and efficiency by which *GFP* transcription and translation are repressed. Western blot analysis will be performed to determine the rate and extent of tagged GFP degradation. Finally, in all cases, biomass and cell viability will be monitored using optical density measurements (OD_{600}) and commercially available cell viability kits, respectively to gain a preliminary insight regarding possible phenotype switch impacts to overall host fitness.

5.1.6. Anticipated Outcomes

The proposed dynamic phenotype switches described above all share the same desired output (Figure 5.8). However, while the complete loss of further GFP expression and detection is desirable, it is possible that the subsequent complete absence of GFP may not be achieved via the implementation of the transcriptional and translational approaches. In these cases, GFP levels are expected to remain similar to those present at the time of phenotype switch induction. It should be noted that, while residual GFP does not impact our test system, residual competing pathway genes (as represented by GFP) would result in an ineffective flux redistribution system with little observable impact on final production metrics, as further discussed in the subsequent sections. To this end, a combination of the approaches described above may ultimately be required to achieve a phenotype switch with the ideal response. For example, CRISPRi and protein degradation are hypothesized to result in a more desirable (ideal) response.

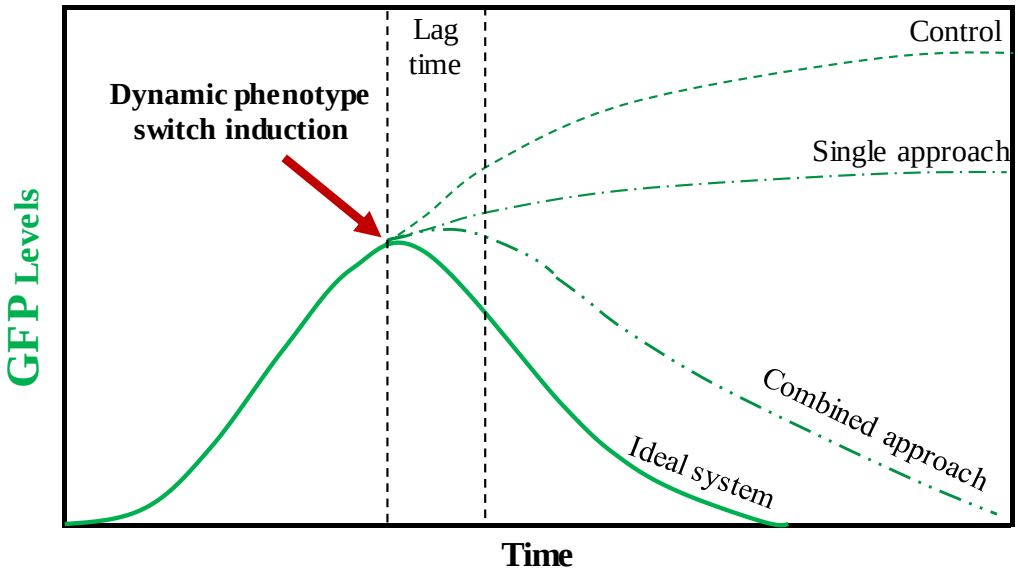


Figure 5.8 Dynamic Phenotype Switch Output. The ideal system (solid line) will result in a rapid and efficient response relative to the control (dash line), observed by the loss of detectable GFP. Most individual techniques will likely result in steady-state GFP levels close to those detected at the time of phenotype switch induction (dash-dot line). However, by combining more than one approach, a response (dash-dot-dot line) mirroring the ideal system is expected.

5.1.7. Pitfalls and Proposed Solutions

Should initial challenges arise while investigating the various phenotype control switches outlined above, initial proposed solutions will be implemented. CRISPRi and sRNA mediated phenotype switches may be leaky and result in undesirable GFP repression prior to system induction. Should this be the case, alternative promoters and/or promoter-repressor systems will be investigated to more tightly regulate expression. Meanwhile, while the clear majority of proteins in *E. coli* are reported to be accessible to *mf*-Lon proteolysis, select proteins have been reported to be stable in the presence of *mf*-Lon protease despite containing the pdt#3 degradation tag. Although the exact mechanism for avoiding degradation is unknown, protein folding and/or post-translational modifications are suspected causes. Should this be the case,

alternative protein degradation strategies, including native *E. coli* proteolysis mechanisms will be investigated. In the case of the dynamic chromosomal deletion approach, less than 100% efficient GFP excision will result in two subpopulations containing and lacking GFP. If necessary, strategies for controlled cell death will be investigated. In this case, a toxin-antitoxin system will be investigated to enable survival only following successful recombination.

5.2. Dynamic Flux Regulation for Phenol and MA Production

Phenol and muconic acid (MA) are important platform chemicals, used as platform chemicals for the synthesis of plastics and polymers. In the case of phenol, the complete biosynthesis from glucose has been reported stemming directly from endogenous chorismate. By eliminating chorismate mutase activity, encoded by native *pheA*, significant phenol accumulation was realized. However, while this strategy increased the available intracellular pool of chorismate, disrupting native *pheA* necessitates phenylalanine supplementation in a minimal salts medium to support cell growth. On the other hand, MA biosynthesis was first reported following the development of a three step pathway stemming from endogenous 3-dehydroshikimate (3DHS). Although several other successful pathways have been engineered instead stemming from endogenous chorismate, the original route confers highest achievable yields. Whereas 3DHS is a key intermediate metabolite in the shikimic acid pathway, significant MA overproduction requires the deletion of subsequent shikimate dehydrogenase, encoded by native *aroE*. However, this strategy results in multiple essential vitamin auxotrophies, including phenylalanine, tyrosine, tryptophan, *p*-aminobenzoate, *p*-hydroxybenzoate, and 2,3-dihydroxybenzoate. Consequently, all essential vitamins must be supplemented in a minimal salts media to support cell viability, ultimately leading to expensive media formulations and poor scalability.

As an alternative approach, the best two or three phenotype switches (as identified by the previous proposed work) to effectively simulate the traditional chromosomal deletion will be implemented for enhanced phenol and MA biosynthesis. Select phenotype switches will be implemented to the original MA pathway to modulate native *aroE* expression and in the case of phenol and the MA funneling pathway combination to modulate native *pheA*. By eliminating *aroE* and *pheA* expression, cells are hypothesized to shift carbon flux away from key nutrient pathways and towards MA and phenol production, resulting in non-growing cells. Cells existing in a non-growing yet metabolically active state can sustain high biosynthetic capacity. For example, the exogenous addition of chloramphenicol (inhibits ribosome function) was previously demonstrated to decouple growth from phenylalanine production. Although this approach resulted in a 50% greater specific productivity over the control strain, the high cost of antibiotics and growing concern of resistant mutants renders this approach impractical for biosynthetic applications.

To evaluate the applicability of the select dynamic phenotype switches, related host backgrounds will be investigated. *E. coli* NST74 Δ*aroE* will serve as the established gene deletion control to evaluate the relative performance of the original 3DHS derived pathway, while *E. coli* NST74 Δ*pheA* will serve as the deletion control to evaluate the relative performance of phenol biosynthesis and the MA 'funneling' combination pathway. Meanwhile, *E. coli* NST74, is a previously engineered phenylalanine overproducer (essential metabolite downstream of *aroE* and *pheA*) and will serve as the control strain for both MA biosynthesis routes. For all the dynamic phenotype switches tested, this host will serve as an ideal test platform because efficient repression of shikimate dehydrogenase activity (encoded by *aroE*) and chorismate mutase (encoded by *pheA*) by any dynamic approach will result in reduced carbon flux toward phenylalanine. Accordingly, phenylalanine titers will be monitored and

compared to the control as one metric to determine effectiveness of each strategy utilized. In the case of sRNA and CRISPRi implementation, RT-PCR and/or qPCR can be used to evaluate relative expression levels and rate and efficiency of repression. Meanwhile, in the case of *mf*-Lon mediated protein degradation, while phenylalanine titers will provide useful insight, western-blot analysis will ultimately be required to confirm the degradation of AroE and PheA. Depending on the outcomes (e.g. relative phenylalanine levels), the same tuning strategies as described previously (i.e. promoter strength, plasmid copy number, RBS, etc.) will be implemented to identify a host capable of achieving the dynamic characteristics closest to the ideal outcome. Finally, cell viability over the course of dynamic control strategy implementation will be evaluated using commercially available kits.

5.3. Conclusions

These works demonstrate the applicability of various dynamic phenotype switches as a control strategy for metabolic flux control. Novel pathway design and construction enabled phenol and MA production while avoiding some of the inherent limitations associated with the original biosynthesis routes. To further enhance production metrics, numerous dynamic flux control strategies, and combinations thereof, should be characterized and utilized as an alternative to traditional gene disruptions.

REFERENCES

Achkar J, Xian M, Zhao H, Frost JW. 2005. Biosynthesis of phloroglucinol. J Am Chem Soc 127(15):5332-3.

Achmon Y, Ben-Barak Zelas Z, Fishman A. 2014. Cloning Rosa hybrid phenylacetaldehyde synthase for the production of 2-phenylethanol in a whole cell Escherichia coli system. Appl Microbiol Biotechnol 98(8):3603-11.

Adkins J, Jordan J, Nielsen DR. 2013. Engineering Escherichia coli for renewable production of the 5-carbon polyamide building-blocks 5-aminovalerate and glutarate. Biotechnol Bioeng 110(6):1726-1734.

Adkins J, Pugh S, McKenna R, Nielsen DR. 2012. Engineering microbial chemical factories to produce renewable "biomonomers". Front Microbiol 3:313.

Akashi H, Gojobori T. 2002. Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. Proc Natl Acad Sci U S A 99(6):3695-700.

Atsumi S, Wu TY, Eckl EM, Hawkins SD, Buelter T, Liao JC. 2010. Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. Appl Microbiol Biotechnol 85(3):651-7.

Averesch NJH, Krömer JO. 2014. Tailoring strain construction strategies for muconic acid production in S. cerevisiae and E. coli. Metabolic Engineering Communications 1:19-28.

Beekrum S, Govinden R, Padayachee T, Odhav B. 2003. Naturally occurring phenols: a detoxification strategy for fumonisin B1. Food Addit Contam 20(5):490-3.

Berry A, Dodge TC, Pepsin M, Weyler W. 2002. Application of metabolic engineering to improve both the production and use of biotech indigo. Journal of Industrial Microbiology & Biotechnology 28(3):127-133.

Birnbaum S, Bailey JE. 1991. Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinant Escherichia coli. Biotechnology and bioengineering 37(8):736-45.

Bongaerts J, Kromer M, MoLler U, Raeven L, Wubbolts M. 2001. Metabolic engineering for microbial production of aromatic amino acids and derived compounds. Metab Eng 3.

Bouwer EJ, Zehnder AJ. 1993. Bioremediation of organic compounds--putting microbial metabolism to work. Trends Biotechnol 11(8):360-7.

Brockman IM, Prather KL. 2015. Dynamic knockdown of E. coli central metabolism for redirecting fluxes of primary metabolites. Metab Eng 28:104-13.

Buss K, Müller R, Dahm C, Gaitatzis N, Skrzypczak-Pietraszek E, Lohmann S, Gassen M, Leistner E. 2001. Clustering of isochorismate synthase genes menF and entC and channeling of isochorismate in Escherichia coli. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1522(3):151-157.

Cafaro V, Izzo V, Scognamiglio R, Notomista E, Capasso P, Casbarra A, Pucci P, Di Donato A. 2004. Phenol hydroxylase and toluene/o-xylene monooxygenase from Pseudomonas stutzeri OX1: interplay between two enzymes. Appl Environ Microbiol 70(4):2211-9.

Cafaro V, Scognamiglio R, Viggiani A, Izzo V, Passaro I, Notomista E, Piaz FD, Amoresano A, Casbarra A, Pucci P and others. 2002. Expression and purification of the recombinant subunits of toluene/o-xylene monooxygenase and reconstitution of the active complex. European journal of biochemistry / FEBS 269(22):5689-99.

Cameron DE, Collins JJ. 2014. Tunable protein degradation in bacteria. Nat Biotech 32(12):1276-1281.

Cao B, Geng A, Loh KC. 2008. Induction of ortho- and meta-cleavage pathways in Pseudomonas in biodegradation of high benzoate concentration: MS identification of catabolic enzymes. Appl Microbiol Biotechnol 81(1):99-107.

Carothers JM, Goler JA, Keasling JD. 2009. Chemical synthesis using synthetic biology. Curr Opin Biotechnol 20(4):498-503.

Caspi R, Foerster H, Fulcher CA, Kaipa P, Krummenacker M, Latendresse M, Paley S, Rhee SY, Shearer AG, Tissier C and others. 2008. The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Research 36:D623-D631.

Castaño-Cerezo S, Pastor JM, Renilla S, Bernal V, Iborra JL, Cánovas M. 2009. An insight into the role of phosphotransacetylase (pta) and the acetate/acetyl-CoA node in Escherichia coli. Microbial Cell Factories 8(1):54.

Cochrane FC, Davin LB, Lewis NG. 2004. The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65(11):1557-64.

Council BRA, editor. 2006. Biofuels in the European Union: A vision for 2030 and beyond. Belgium: Biofuels Research Advisory Council, European Commission.

Curran KA, Alper HS. 2012. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. Metab Eng 14(4):289-97.

Curran KA, Leavitt JM, Karim AS, Alper HS. 2013. Metabolic engineering of muconic acid production in Saccharomyces cerevisiae. Metab Eng 15.

Dahm C, Müller R, Schulte G, Schmidt K, Leistner E. 1998. The role of isochorismate hydroxymutase genes entC and menF in enterobactin and menaquinone biosynthesis in Escherichia coli. Biochimica et Biophysica Acta (BBA) - General Subjects 1425(2):377-386.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640-5.

Deng Y, Ma LZ, Mao Y. 2016. Biological production of adipic acid from renewable substrates: Current and future methods. Biochemical Engineering Journal 105:16-26.

Diaz E. 2004. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int Microbiol 7(3):173-80.

Douglas CJ. 1996. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. Trends in Plant Science 1(6):171-178.

Draths KM, Frost JW. 1994. Environmentally Compatible Synthesis of Adipic Acid from D-Glucose. Journal of the American Chemical Society 116(1):399-400.

Floras N, Xiao J, Berry A, Bolivar F, Valle F. 1996. Pathway engineering for the production of aromatic compounds in Escherichia coli. Nat Biotech 14(5):620-623.

Foor F, Morin N, Bostian KA. 1993. Production of L-dihydroxyphenylalanine in Escherichia coli with the tyrosine phenol-lyase gene cloned from Erwinia herbicola. Appl Environ Microbiol 59(9):3070-5.

Fullenkamp DE, Barrett DG, Miller DR, Kurutz JW, Messersmith PB. 2014. pH-dependent cross-linking of catechols through oxidation via Fe(3+) and potential implications for mussel adhesion. RSC Advances 4(48):25127-25134.

Gaille C, Kast P, Haas D. 2002. Salicylate biosynthesis in Pseudomonas aeruginosa. Purification and characterization of PchB, a novel bifunctional enzyme displaying isochorismate pyruvate-lyase and chorismate mutase activities. J Biol Chem 277(24):21768-75.

Gallagher DT, Mayhew M, Holden MJ, Howard A, Kim KJ, Vilker VL. 2001. The crystal structure of chorismate lyase shows a new fold and a tightly retained product. Proteins 44(3):304-11.

Gibson D. 2009. One-step enzymatic assembly of DNA molecules up to several hundred kilobases in size.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth 6(5):343-345.

Gosset G. 2005. Improvement of Escherichia coli production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. Microbial Cell Factories 4(1):14.

Gosset G. 2009. Production of aromatic compounds in bacteria. Curr Opin Biotechnol 20.

Gottlieb K, Albermann C, Sprenger GA. 2014. Improvement of L-phenylalanine production from glycerol by recombinant Escherichia coli strains: the role of extra copies of glpK, glpX, and tktA genes. Microb Cell Fact 13(1):96.

Gu P, Yang F, Wang Q, Huang J, Qi Q. 2012. One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in Escherichia coli. Microbial Cell Factories.

Guest JR. 1977. Menaquinone biosynthesis: mutants of Escherichia coli K-12 requiring 2-succinylbenzoate. Journal of Bacteriology 130(3):1038-1046.

Hawkins JS, Wong S, Peters JM, Almeida R, Qi LS. 2015. Targeted Transcriptional Repression in Bacteria Using CRISPR Interference (CRISPRi). Methods in molecular biology (Clifton, N.J.) 1311:349-62.

Henry CS, Broadbelt LJ, Hatzimanikatis V. 2010. Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate. Biotechnology and bioengineering 106(3):462-473.

Herrmann K. 1995. The Shikimate Pathway as an Entry to Aromatic Secondary Metabolism. Plant Physiol 107:7-12.

Isken S, de Bont JAM. 1998. Bacteria tolerant to organic solvents. Extremophiles 2(3):229-238.

Iwasaki Y, Gunji H, Kino K, Hattori T, Ishii Y, Kirimura K. 2010. Novel metabolic pathway for salicylate biodegradation via phenol in yeast Trichosporon moniliiforme. Biodegradation 21(4):557-64.

Jiang X, Meng X, Xian M. 2009. Biosynthetic pathways for 3-hydroxypropionic acid production. Appl Microbiol Biotechnol 82(6):995-1003.

Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. 2015. Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. Appl Environ Microbiol 81(7):2506-14.

Jindrová E, Chocová M, Demnerová K, Brenner V. 2002. Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiologica 47(2):83-93.

Johnson CW, Salvachúa D, Khanna P, Smith H, Peterson DJ, Beckham GT. 2016. Enhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity. Metabolic Engineering Communications 3:111-119.

Juminaga D, Baidoo EE, Redding-Johanson AM, Batth TS, Burd H, Mukhopadhyay A, Petzold CJ, Keasling JD. 2012. Modular engineering of L-tyrosine production in Escherichia coli. Appl Environ Microbiol 78(1):89-98.

Kaleta C, Schauble S, Rinas U, Schuster S. 2013. Metabolic costs of amino acid and protein production in Escherichia coli. Biotechnol J 8(9):1105-14.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27-30.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42(Database issue):D199-205.

Kang SY, Choi O, Lee JK, Hwang BY, Uhm TB, Hong YS. 2012. Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing Escherichia coli strain. Microb Cell Fact 11:153.

Kang Z, Zhang C, Du G, Chen J. 2014. Metabolic engineering of Escherichia coli for production of 2-phenylethanol from renewable glucose. Appl Biochem Biotechnol 172(4):2012-21.

Kikuchi Y, Tsujimoto K, Kurahashi O. 1997. Mutational analysis of the feedback sites of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase of Escherichia coli. Appl Environ Microbiol 63(2):761-762.

Kim B, Park H, Na D, Lee SY. 2014. Metabolic engineering of Escherichia coli for the production of phenol from glucose. Biotechnology Journal 9(5):621-629.

Kirimura K, Gunji H, Wakayama R, Hattori T, Ishii Y. 2010. Enzymatic Kolbe-Schmitt reaction to form salicylic acid from phenol: enzymatic characterization and gene identification of a novel enzyme, Trichosporon moniliiforme salicylic acid decarboxylase. Biochem Biophys Res Commun 394(2):279-84.

Koma D, Yamanaka H, Moriyoshi K, Ohmoto T, Sakai K. 2012. Production of aromatic compounds by metabolically engineered Escherichia coli with an expanded shikimate pathway. Appl Environ Microbiol 78(17):6203-16.

Kromer JO, Wittmann C, Schroder H, Heinzle E. 2006. Metabolic pathway analysis for rational design of L-methionine production by Escherichia coli and Corynebacterium glutamicum. Metab Eng 8(4):353-69.

Kukor JJ, Olsen RH, Ballou DP. 1988. Cloning and expression of the catA and catBC gene clusters from Pseudomonas aeruginosa PAO. Journal of Bacteriology 170(10):4458-4465.

Kumar V, Ashok S, Park S. 2013. Recent advances in biological production of 3-hydroxypropionic acid. Biotechnology Advances 31(6):945-961.

Kunjapur AM, Tarasova Y, Prather KL. 2014. Synthesis and accumulation of aromatic aldehydes in an engineered strain of Escherichia coli. J Am Chem Soc 136(33):11644-54.

Kwon O, Hudspeth ME, Meganathan R. 1996. Anaerobic biosynthesis of enterobactin Escherichia coli: regulation of entC gene expression and evidence against its involvement in menaquinone (vitamin K2) biosynthesis. Journal of Bacteriology 178(11):3252-3259.

Lamb AL. 2011. Pericyclic reactions catalyzed by chorismate-utilizing enzymes. Biochemistry 50(35):7476-7483.

Lee JW, Kim HU, Choi S, Yi J, Lee SY. 2011. Microbial production of building block chemicals and polymers. Curr Opin Biotechnol 22(6):758-67.

Liao H, Gokarn R, Gort S, Jessen H, Selifonova O. 2007. Production of 3-hydroxypropionic acid using beta-alanine/pyruvate aminotransferase. US Patent.

Lin Y, Shen X, Yuan Q, Yan Y. 2013. Microbial biosynthesis of the anticoagulant precursor 4-hydroxycoumarin. Nat Commun 4.

Lin Y, Sun X, Yuan Q, Yan Y. 2014. Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in Escherichia coli. Metab Eng 23.

Lin Y, Yan Y. 2012. Biosynthesis of caffeic acid in Escherichia coli using its endogenous hydroxylase complex. Microb Cell Fact 11:42.

Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT and others. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proc Natl Acad Sci U S A 111(33):12013-8.

Litcofsky KD, Afeyan RB, Krom RJ, Khalil AS, Collins JJ. 2012. Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. Nature methods 9(11):1077-80.

Liu L, Duan X, Wu J. 2016. L-Tryptophan Production in Escherichia coli Improved by Weakening the Pta-AckA Pathway. Plos One 11(6):e0158200.

Liu SP, Liu RX, Xiao MR, Zhang L, Ding ZY, Gu ZH, Shi GY. 2014. A systems level engineered E. coli capable of efficiently producing L-phenylalanine. Process Biochemistry 49(5):751-757.

Lupa B, Lyon D, Gibbs MD, Reeves RA, Wiegel J. 2005. Distribution of genes encoding the microbial non-oxidative reversible. Genomics 86(3).

Lütke-Eversloh T, Stephanopoulos G. 2005. Feedback Inhibition of Chorismate Mutase/Prephenate Dehydrogenase (TyrA) of Escherichia coli: Generation and Characterization of Tyrosine-Insensitive Mutants. Appl Environ Microbiol 71(11):7224-7228.

Mao Z, Shin HD, Chen R. 2009. A recombinant E. coli bioprocess for hyaluronan synthesis. Appl Microbiol Biotechnol 84(1):63-9.

Martin CH, Nielsen DR, Solomon KV, Prather KL. 2009. Synthetic metabolism: engineering biology at the protein and pathway scales. Chemistry & biology 16(3):277-86.

Masuo S, Osada L, Zhou S, Fujita T, Takaya N. 2015. Aspergillus oryzae pathways that convert phenylalanine into the flavor volatile 2-phenylethanol. Fungal Genet Biol 77:22-30.

Matsui T, Yoshida T, Hayashi T, Nagasawa T. 2006. Purification, characterization, and gene cloning of 4-hydroxybenzoate decarboxylase of Enterobacter cloacae P240. Arch Microbiol 186(1):21-9.

McKenna R, Nielsen DR. 2011. Styrene biosynthesis from glucose by engineered E. coli. Metab Eng 13(5):544-54.

McKenna R, Pugh S, Thompson B, Nielsen DR. 2013. Microbial Production of the Aromatic Building-Blocks (S)-Styrene Oxide and (R)-1,2-Phenylethanediol from Renewable Resources. Biotechnol. J. 8(12):1465-1475.

McKenna R, Thompson B, Pugh S, Nielsen DR. 2014. Rational and combinatorial approaches to engineering styrene production by Saccharomyces cerevisiae. Microb Cell Fact 13:123.

Miao L, Li Q, Diao A, Zhang X, Ma Y. 2015a. Construction of a novel phenol synthetic pathway in Escherichia coli through 4-hydroxybenzoate decarboxylation. Appl Microbiol Biotechnol.

Miao L, Li Q, Diao A, Zhang X, Ma Y. 2015b. Construction of a novel phenol synthetic pathway in Escherichia coli through 4-hydroxybenzoate decarboxylation. Appl Microbiol Biotechnol 99(12):5163-73.

Myers RL. 2007. The 100 most important chemical compounds : a reference guide. Westport, Conn.: Greenwood Press. xxvi, 326 p. p.

Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY. 2013. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat Biotech 31(2):170-174.

Nagasawa T, Utagawa T, Goto J, Kim CJ, Tani Y, Kumagai H, Yamada H. 1981. Syntheses of L-tyrosine-related amino acids by tyrosine phenol-lyase of Citrobacter intermedius. Eur J Biochem 117(1):33-40.

Nexant I. 2011. Petrochemical Market Dynamics - Aromatics (PPE11-G-PCMD-Aromatics).

Nikolau BJ, Perera MADN, Brachova L, Shanks B. 2008. Platform biochemicals for a biorenewable chemical industry. Plant Journal 54(4):536-545.

Niu W, Draths KM, Frost JW. 2002. Benzene-free synthesis of adipic acid. Biotechnol Prog 18(2):201-11.

Noda S, Shirai T, Oyama S, Kondo A. 2016. Metabolic design of a platform Escherichia coli strain producing various chorismate derivatives. Metab Eng 33:119-129.

Notomista E, Lahm A, Di Donato A, Tramontano A. 2003. Evolution of bacterial and archaeal multicomponent monooxygenases. J Mol Evol 56(4):435-45.

Para G, Lucciardi P, Baratti J. 1985. Synthesis of L-Tyrosine by Immobilized Escherichia-Intermedia Cells. Applied Microbiology and Biotechnology 21(5):273-279.

Pérez-Gil J, Rodríguez-Concepción M. 2013. Metabolic plasticity for isoprenoid biosynthesis in bacteria. Biochemical Journal 452(1):19-25.

Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC, editors. 2005. Biomass as Feedstock for a bioenergy and bioproducts Industry: the technical feasibility of a billion-ton annual supply. Oak Ridge, TN: U.S. Department of Energy.

Pittard J, Camakaris H, Yang J. 2005. The TyrR regulon. Molecular microbiology 55(1):16-26.

Postma PW, Lengeler JW, Jacobson GR. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol Rev 57.

Pugh S, McKenna R, Osman M, Thompson B, Nielsen DR. 2014. Rational engineering of a novel pathway for producing the aromatic compounds p-hydroxybenzoate, protocatechuate, and catechol in Escherichia coli. Process Biochemistry 49(11):1843-1850.

Qi WW, Vannelli T, Breinig S, Ben-Bassat A, Gatenby AA, Haynie SL, Sariaslani FS. 2007. Functional expression of prokaryotic and eukaryotic genes in Escherichia coli for conversion of glucose to p-hydroxystyrene. Metab Eng 9(3):268-76.

Quan J, Tian J. 2011. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. Nat. Protocols 6(2):242-251.

Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzalez MI, Rojas A, Teran W, Segura A. 2002. Mechanisms of solvent tolerance in gram-negative bacteria. Annual Review of Microbiology 56:743-768.

Reed JL, Vo TD, Schilling CH, Palsson BO. 2003. An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). Genome Biol 4(9):R54.

Ren YX, Yang S, Yuan QP, Sun XX. 2015. Microbial production of phenol via salicylate decarboxylation. Rsc Advances 5(112):92685-92689.

Rodriguez A, Martínez JA, Flores N, Escalante A, Gosset G, Bolivar F. 2014. Engineering Escherichia coli to overproduce aromatic amino acids and derived compounds. Microbial Cell Factories 13.

Santos CN, Xiao W, Stephanopoulos G. 2012. Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in Escherichia coli. Proc Natl Acad Sci U S A 109(34):13538-43.

Satoh Y, Tajima K, Munekata M, Keasling JD, Lee TS. 2012a. Engineering of a tyrosol-producing pathway, utilizing simple sugar and the central metabolic tyrosine, in Escherichia coli. J Agric Food Chem 60(4):979-84.

Satoh Y, Tajima K, Munekata M, Keasling JD, Lee TS. 2012b. Engineering of L-tyrosine oxidation in Escherichia coli and microbial production of hydroxytyrosol. Metab Eng 14(6):603-10.

Sengupta S, Jonnalagadda S, Goonewardena L, Juturu V. 2015. Metabolic Engineering of a Novel Muconic Acid Biosynthesis Pathway via 4-Hydroxybenzoic Acid in Escherichia coli. Appl Environ Microbiol 81(23):8037-8043.

Shen CR, Lan EI, Dekishima Y, Baez A, Cho KM, Liao JC. 2011. Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli. Appl Environ Microbiol 77(9):2905-2915.

Shiloach J, Kaufman J, Guillard AS, Fass R. 1996. Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by Escherichia coli BL21 (lambdaDE3) and Escherichia coli JM109. Biotechnology and bioengineering 49(4):421-8.

Shiue E, Prather KL. 2014. Improving D-glucaric acid production from myo-inositol in E. coli by increasing MIOX stability and myo-inositol transport. Metab Eng 22:22-31.

Siebert M, Severin K, Heide L. 1994. Formation of 4-hydroxybenzoate in Escherichia coli: characterization of the ubiC gene and its encoded enzyme chorismate pyruvate-lyase. Microbiology (Reading, England) 140 (Pt 4):897-904.

Sonoki T, Morooka M, Sakamoto K, Otsuka Y, Nakamura M, Jellison J, Goodell B. 2014. Enhancement of protocatechuate decarboxylase activity for the effective production of muconate from lignin-related aromatic compounds. J Biotechnol 192 Pt A:71-7.

Sridevi V, Chandana Lakshami M, Manasa M, Sravani M. 2012. Metabolic pathways for the biodegradation of phenol. Int J Eng Sci Adv Technol 2:695-705.

Stratford M, Plumridge A, Archer DB. 2007. Decarboxylation of sorbic acid by spoilage yeasts is associated with the PAD1 gene. Applied and Environmental Microbiology 73(20):6534-6542.

Sun X, Lin Y, Huang Q, Yuan Q, Yan Y. 2013. A novel muconic acid biosynthesis approach by shunting tryptophan biosynthesis via anthranilate. Appl Environ Microbiol 79.

Sun X, Lin Y, Yuan Q, Yan Y. 2014. Biological production of muconic acid via a prokaryotic 2,3-dihydroxybenzoic acid decarboxylase. ChemSusChem 7(9):2478-81.

Sun Z, Ning Y, Liu L, Liu Y, Sun B, Jiang W, Yang C, Yang S. 2011. Metabolic engineering of the L-phenylalanine pathway in Escherichia coli for the production of S- or R-mandelic acid. Microb Cell Fact 10:71.

Tao Y, Fishman A, Bentley WE, Wood TK. 2004. Oxidation of benzene to phenol, catechol, and 1,2,3-trihydroxybenzene by toluene 4-monooxygenase of Pseudomonas mendocina KR1 and toluene 3-monooxygenase of Ralstonia pickettii PKO1. Appl Environ Microbiol 70(7):3814-20.

Terzer M, Stelling J. 2008. Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 24(19):2229-35.

Thompson B, Machas M, Nielsen DR. 2016. Engineering and comparison of non-natural pathways for microbial phenol production. Biotechnology and bioengineering.

Tinberg CE, Song WJ, Izzo V, Lippard SJ. 2011. Multiple Roles of Component Proteins in Bacterial Multicomponent Monooxygenases: Phenol Hydroxylase and Toluene/o-Xylene Monooxygenase from Pseudomonas sp. OX1. Biochemistry 50(11):1788-1798.

Tran NP, Gury J, Dartois V, Nguyen TK, Seraut H, Barthelmebs L, Gervais P, Cavin JF. 2008. Phenolic acid-mediated regulation of the padC gene, encoding the phenolic acid decarboxylase of Bacillus subtilis. J Bacteriol 190(9):3213-24.

Tribe DE. 1987. Novel microorganism and method. US Patent 4,681,852.

Tsuruno K, Honjo H, Hanai T. 2015. Enhancement of 3-hydroxypropionic acid production from glycerol by using a metabolic toggle switch. Microbial Cell Factories 14:155.

van Schie PM, Young LY. 2000. Biodegradation of Phenol: Mechanisms and Applications. Bioremediation Journal 4(1):1-18.

Vardar G, Wood TK. 2005. Alpha-Subunit Positions Methionine 180 and Glutamate 214 of Pseudomonas stutzeri OX1 Toluene-o-Xylene Monooxygenase Influence Catalysis. Journal of Bacteriology 187(4):1511-1514.

Vargas-Tah A, Martinez LM, Hernandez-Chavez G, Rocha M, Martinez A, Bolivar F, Gosset G. 2015. Production of cinnamic and p-hydroxycinnamic acid from sugar mixtures with engineered Escherichia coli. Microb Cell Fact 14(1):6.

Wang J, Cheng LK, Wang J, Liu Q, Shen T, Chen N. 2013. Genetic engineering of Escherichia coli to enhance production of L-tryptophan. Appl Microbiol Biotechnol 97(17):7587-96.

Wang S, Zhang S, Xiao A, Rasmussen M, Skidmore C, Zhan J. 2015. Metabolic engineering of Escherichia coli for the biosynthesis of various phenylpropanoid derivatives. Metab Eng 29:153-159.

Watts KT, Mijts BN, Lee PC, Manning AJ, Schmidt-Dannert C. 2006. Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. Chem Biol 13(12):1317-26.

Weber C, Brückner C, Weinreb S, Lehr C, Essl C, Boles E. 2012. Biosynthesis of cis, cis-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by Saccharomyces cerevisiae. Appl Environ Microbiol 78.

Weber M, Weber M, Kleine-Boymann M. 2000. Phenol. Ullmann's Encyclopedia of Industrial Chemistry: Wiley-VCH Verlag GmbH & Co. KGaA.

Weikert C, Sauer U, Bailey JE. 1998. Increased phenylalanine production by growing and nongrowing Escherichia coli strain CWML2. Biotechnol Prog 14(3):420-4.

Wierckx NJ, Ballerstedt H, de Bont JA, de Winde JH, Ruijssenaars HJ, Wery J. 2008. Transcriptome analysis of a phenol-producing Pseudomonas putida S12 construct: genetic and physiological basis for improved production. J Bacteriol 190(8):2822-30.

Wierckx NJ, Ballerstedt H, de Bont JA, Wery J. 2005. Engineering of solvent-tolerant Pseudomonas putida S12 for bioproduction of phenol from glucose. Appl Environ Microbiol 71(12):8221-7.

Williams TC, Averesch NJ, Winter G, Plan MR, Vickers CE, Nielsen LK, Kromer JO. 2015. Quorum-sensing linked RNA interference for dynamic metabolic pathway control in Saccharomyces cerevisiae. Metab Eng 29:124-34.

Wolfe AJ. 2005. The Acetate Switch. Microbiology and Molecular Biology Reviews 69(1):12-50.

Wu JJ, Du GC, Zhou JW, Chen J. 2013. Metabolic engineering of Escherichia coli for (2S)-pinocembrin production from glucose by a modular metabolic strategy. Metabolic Engineering 16:48-55.

Xie D, Shao Z, Achkar J, Zha W, Frost JW, Zhao H. 2006. Microbial synthesis of triacetic acid lactone. Biotechnology and bioengineering 93(4):727-36.

Yang J, Guo L. 2014. Biosynthesis of β-carotene in engineered E. coli using the MEP and MVA pathways. Microbial Cell Factories 13:160.

Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R and others. 2011. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. Nature Chemical Biology 7(7):445-452.

Youngquist JT, Rose JP, Pfleger BF. 2013. Free fatty acid production in Escherichia coli under phosphate-limited conditions. Appl Microbiol Biotechnol 97(11):5149-59.

Zhang H, Li Z, Pereira B, Stephanopoulos G. 2015a. Engineering E. coli–E. coli cocultures for production of muconic acid from glycerol. Microbial Cell Factories 14(1):134.

Zhang H, Pereira B, Li Z, Stephanopoulos G. 2015b. Engineering Escherichia coli coculture systems for the production of biochemical products. Proc Natl Acad Sci USA 112.

Zhang H, Pereira B, Li Z, Stephanopoulos G. 2015c. Engineering Escherichia coli coculture systems for the production of biochemical products. Proc Natl Acad Sci U S A 112(27):8266-71.

Zhang H, Stephanopoulos G. 2013. Engineering E. coli for caffeic acid biosynthesis from renewable sugars. Appl Microbiol Biotechnol 97(8):3333-41.

Zhang K, Woodruff AP, Xiong M, Zhou J, Dhande YK. 2011. A synthetic metabolic pathway for production of the platform chemical isobutyric acid. ChemSusChem 4(8):1068-70.

Zhao YR, Yang JM, Qin B, Li YH, Sun YZ, Su SZ, Xian M. 2011. Biosynthesis of isoprene in Escherichia coli via methylerythritol phosphate (MEP) pathway. Applied Microbiology and Biotechnology 90(6):1915-1922.

APPENDIX A

SUPPLEMENTAL INFORMATION TO ENGINEERING AND COMPARISON OF NON-NATURAL PATHWAYS FOR MICROBIAL PHENOL PRODUCTION

Table A1 List of primers designed and used for phenol production

Primer	Sequence (5' → 3')
tutA -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GAA TTA TCC GGC AGA ACC
tutA -- XbaI - R	ATT TCT AGA TTA GAT ATA GTC AAA GCG TGC AGT A
pchB -- XbaI - F	AAT ATC TAG ATT CCC GAG AGG TTG CAT GAT GAA AAC T
pchB -- BamHI - R	ATT GGA TCC TTA TGC GGC ACC CCG TGT CTG G
entC -- BamHI - F	ATA GGA TCC AGG AGG ATA AAT AAT GGA TAC GTC ACT GGC TGA
entC -- PstI - R	ATT CTG CAG TTA ATG CAA TCC AAA AAC GTT
SDC -- NcoI - F	ATA CCA TGG AGG AGG ATA AAT AAT GCG TGG TAA AGT TAG CCT G
SDC -- BamHI - R	ATT GGA TCC TTA GGC TTC GCT GTC ATA GAA T
ubiC -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GTC ACA CCC CGC GTT AAC G
ubiC -- BglII - R	ATT AGA TCT TTA GTA CAA CGG TGA CGC CGG TAA A
kpdBCD -- BamHI - F	ATA GGA TCC CCC GTC CGG AGA GGG TAA TTT AAA TAT AAA GTT CG
kpdBCD -- HindIII - R	ATT AAG CTT CTT AGC GGG CCC CTT TAT TAA CGC T
bsdBCD -- NcoI - F	ATA CCA TGG CAG GAG TAT GAT TGA AAT GAA AGC AGA ATT CAA GC
bsdBCD -- PstI - R	ATT CTG CAG AAA GAG ACG GTC TGA CAT ACA GTT CAC GC
aroC-aroA-aroL [pY3] -- F	AGA TCT AAA GGA GGC CAT CCA TGG CTG GAA ACA CAA TTG G
aroC-aroA-aroL [pY3] -- R	ATG CCT GGA GAT CCT TAC TCG AGT TTG GAT CCT C
pY3 [aroC-aroA-aroL] -- F	GAG GAT CCA AAC TCG AGT AAG GAT CTC CAG GCA T
pY3 [aroC-aroA-aroL] -- R	CCA ATT GTG TTT CCA GCC ATG GAT GGC CTC CTT TAG ATC T

Table A2 Comparing maximum product and biomass yields (reported as the fraction of carbon in the glucose substrate converted into phenol product) for each of the three engineered phenol pathways, as predicted using the wild-type *E. coli* network.

Pathway	EFM Count	Max Product Yield	Max Product Yield w/ Growth	Max Biomass Yield
Pathway 1	657,578	0.7448	0.6919	0.7086
Pathway 2	481,279	0.7500	0.6755	0.7086
Pathway 3	481,279	0.7500	0.6755	0.7086

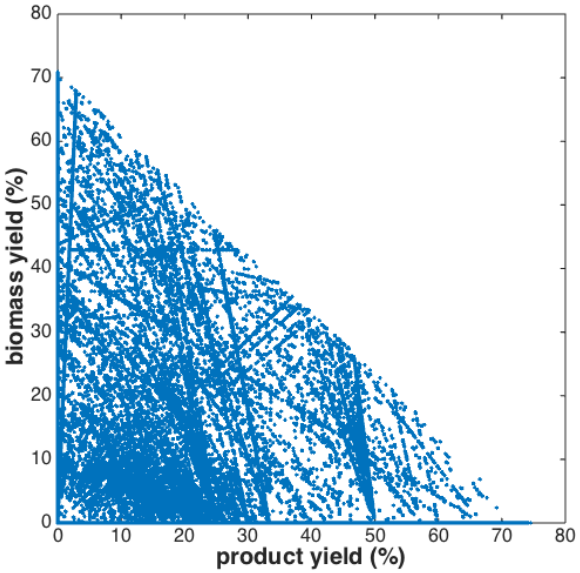


Figure A1a Product vs. biomass yield plot of the EFM distribution of synthesized via the tyrosine route, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

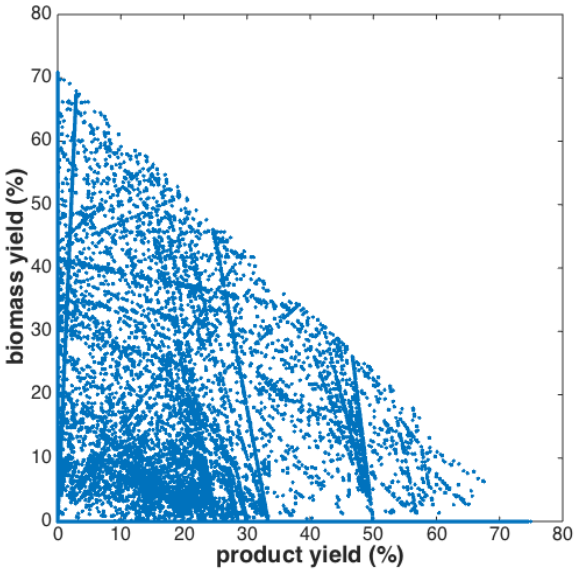


Figure A1b Product vs. biomass yield plot of the EFM distribution of synthesized via the p-hydroxybenzoate route, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

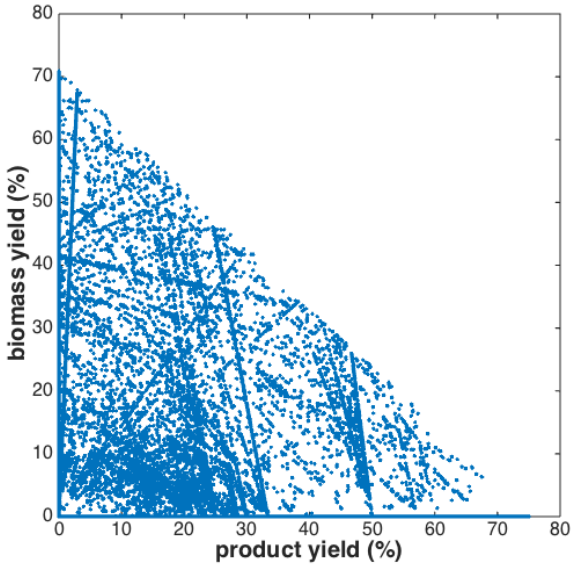


Figure A1c Product vs. biomass yield plot of the EFM distribution of synthesized via the salicylate route, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

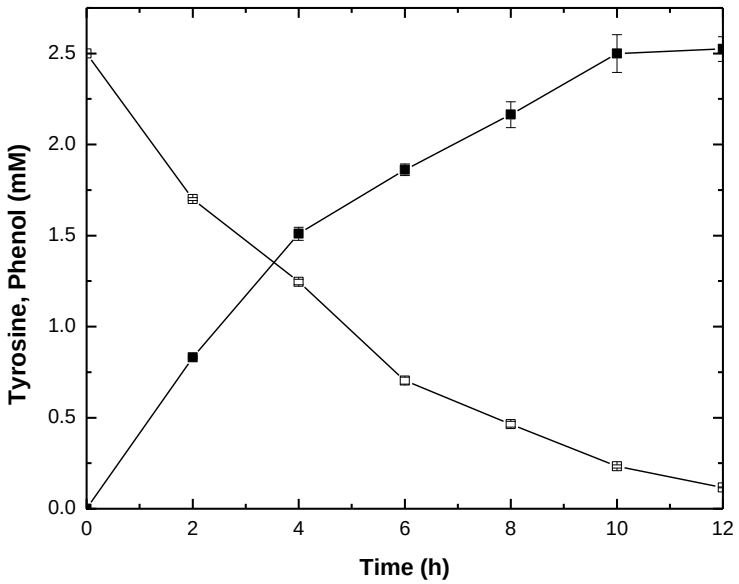


Figure A2 Conversion of tyrosine (open square) to phenol (solid square) by *E. coli* BW25113 pTutA whole resting cells. Error bars represent one standard deviation from triplicate experiment.

Tyrosine route

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R10		'--> P_ex'	Phosphate in
R11		'--> O2'	Oxygen uptake
R12		'--> NH3'	Ammonia uptake
R13		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R14		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R15		'Phenol + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R16		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R17		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R18		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R19		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R20		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R21		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R22		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R23		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R24		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R25		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R26		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R27		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R28		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R29		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R30		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R31		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R32		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R33		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R34		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R35		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R36		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R37		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R38		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R39		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R40		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R41		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R42		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R43		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R44		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R45		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R46		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R47		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R48		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R49		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R50		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R51		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R52		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R53		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R54		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R55		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R56		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R57		'Formate --> CO2'	Formate hydrogenlyase
R58		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R59		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R60		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R61		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R62		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R63		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R64		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R65		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R66		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R67		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R68		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R69		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R70		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R71		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R72		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R73		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R74		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R75		'ATP --> ADP + P'	ATP hydrolysis
R76		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R77		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R78		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R79		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84a		'Chorismate --> Prephenate'	Chorismate mutase (EC 5.4.99.5)
R85a		'Prephenate + NAD --> 4-Hydroxyphenylpyruvate + NADH + CO2'	Prephenate dehydrogenase (EC 1.3.1.12)
R86a		'4-Hydroxyphenylpyruvate + GLU --> 2-OXO + Tyrosine'	Tyrosine aminotransferase (EC 2.6.1.5)
R87a		'Tyrosine <--> PYR + Phenol'	Tyrosine pyruvate lyase (EC 4.1.99.2)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure A3b Networks and functions used for elementary flux mode analysis of phenol biosynthesis via tyrosine.

Salicylate route

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R10		'--> P_ex'	Phosphate in
R11		'--> O2'	Oxygen uptake
R12		'--> NH3'	Ammonia uptake
R13		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R14		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R15		'Phenol + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R16		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R17		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R18		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R19		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R20		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R21		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R22		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R23		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R24		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R25		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R26		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R27		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R28		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R29		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R30		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R31		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R32		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R33		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R34		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R35		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R36		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R37		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R38		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R39		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R40		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R41		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R42		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R43		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R44		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R45		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R46		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R47		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R48		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R49		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R50		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R51		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R52		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R53		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R54		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R55		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R56		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R57		'Formate --> CO2'	Formate hydrogenlyase
R58		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R59		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R60		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R61		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R62		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R63		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R64		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R65		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R66		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R67		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R68		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R69		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R70		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R71		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R72		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R73		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R74		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R75		'ATP --> ADP + P'	ATP hydrolysis
R76		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R77		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R78		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R79		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84b		'Chorismate --> Isochorismate'	Salicylate biosynthesis isochorismate synthase (EC 5.4.4.2)
R85b		'Isochorismate --> Salicylate + PYR'	Isochorismate pyruvavte lyase (EC 4.2.99.21)
R86b		'Salicylate --> Phenol + CO2'	Salicylate decarboxylase (EC 4.1.1.91)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure A3b Networks and functions used for elementary flux mode analysis of phenol biosynthesis via salicylate.

p-Hydroxybenzoate route

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R10		'--> P_ex'	Phosphate in
R11		'--> O2'	Oxygen uptake
R12		'--> NH3'	Ammonia uptake
R13		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R14		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R15		'Phenol + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R16		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R17		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R18		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R19		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R20		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R21		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R22		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R23		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R24		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R25		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R26		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R27		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R28		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R29		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R30		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R31		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R32		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R33		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R34		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R35		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R36		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R37		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R38		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R39		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R40		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R41		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R42		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R43		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R44		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R45		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R46		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R47		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R48		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R49		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R50		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R51		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R52		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R53		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R54		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R55		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R56		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R57		'Formate --> CO2'	Formate hydrogenlyase
R58		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R59		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R60		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R61		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R62		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R63		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R64		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R65		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R66		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R67		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R68		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R69		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R70		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R71		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R72		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R73		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R74		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R75		'ATP --> ADP + P'	ATP hydrolysis
R76		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R77		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R78		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R79		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84c		'Chorismate --> 4-Hydroxybenzoate + PYR'	Chorismate lyase (EC 4.1.3.40)
R85c		'4-Hydroxybenzoate --> Phenol + CO2'	4-hydroxybenzoate decarboxylase (4.1.1.61)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure A3c Networks and functions used for elementary flux mode analysis of phenol biosynthesis via *p*-hydroxybenzoate.

APPENDIX B

SUPPLEMENTAL INFORMATION TO ENGINEERING NOVEL PATHWAYS FOR THE BIOSYNTHESIS OF MUCONIC ACID

Table B1 List of primers designed and used for catechol and muconic acid production

Primer	Sequence (5' → 3')
qsuB -- BamHI - F	ATA GGA TCC AGG AGG ATA AAT AAT GCG TAC ATC CAT TGC CAC TGT TTG
qsuB -- HindIII - R	ATT AAG CTT CTA GTT TGG GAT TCC CCG CTC GA
catA – BglII -F	ATA AGA TCT AGG AGG ATA AAT AAT GAC CGT GAA AAT TTC CCA C
catA -- XbaI - R	ATT TCT AGA TCA GCC CTC CTG CAA CGC
tutA -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GAA TTA TCC GGC AGA ACC
tutA -- XbaI - R	ATT TCT AGA TTA GAT ATA GTC AAA GCG TGC AGT A
SDC -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GCG TGG TAA AGT TAG CCT G
SDC - BamHI - R	ATT GGA TCC TTA GGC TTC GCT GTC ATA GAA T
pchB -- XbaI - F	AAT ATC TAG ATT CCC GAG AGG TTG CAT GAT GAA AAC T
pchB -- BamHI - R	ATT GGA TCC TTA TGC GGC ACC CCG TGT CTG G
entC -- BamHI - F	ATA GGA TCC AGG AGG ATA AAT AAT GGA TAC GTC ACT GGC TGA
entC -- PstI - R	ATT CTG CAG TTA ATG CAA TCC AAA AAC GTT
ubiC -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GTC ACA CCC CGC GTT AAC G
ubiC -- BglII - R	ATT AGA TCT TTA GTA CAA CGG TGA CGC CGG TAA A
kpdBCD -- BamHI - F	ATA GGA TCC CCC GTC CGG AGA GGG TAA TTT AAA TAT AAA GTT CG
kpdBCD -- HindIII - R	ATT AAG CTT CTT AGC GGG CCC CTT TAT TAA CGC T
phKLMNOP -- XbaI - F	ATA TCT AGA AGG AGG ATA AAT AGA GCT CGT GCT GCC TCA CGA
phKLMNOP -- SbfI - R	ATT CCT GCA GGA TGC CCA TGA CTA TAT CTT CTT GAA CAG GGC
pheA (Keio) -- F	CGT GTG AAA CAG AAT GCG AAG ACG AAC AAT A
pheA (Keio) -- R	TAA TCC AGT GCC GGA TGA TTC ACA TCA TC

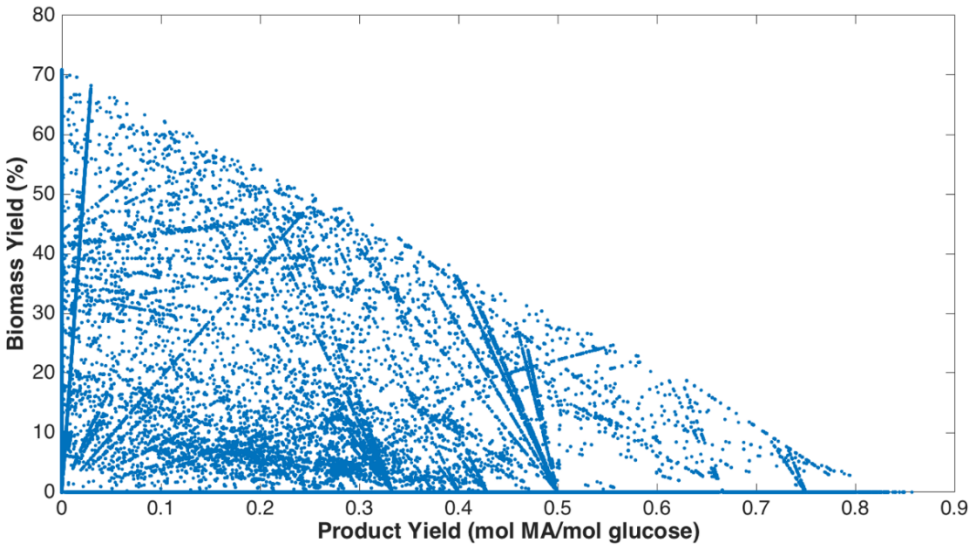


Figure B1a. Product vs. biomass yield plot of the EFM distribution of muconic acid synthesized via the pathway MA1, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

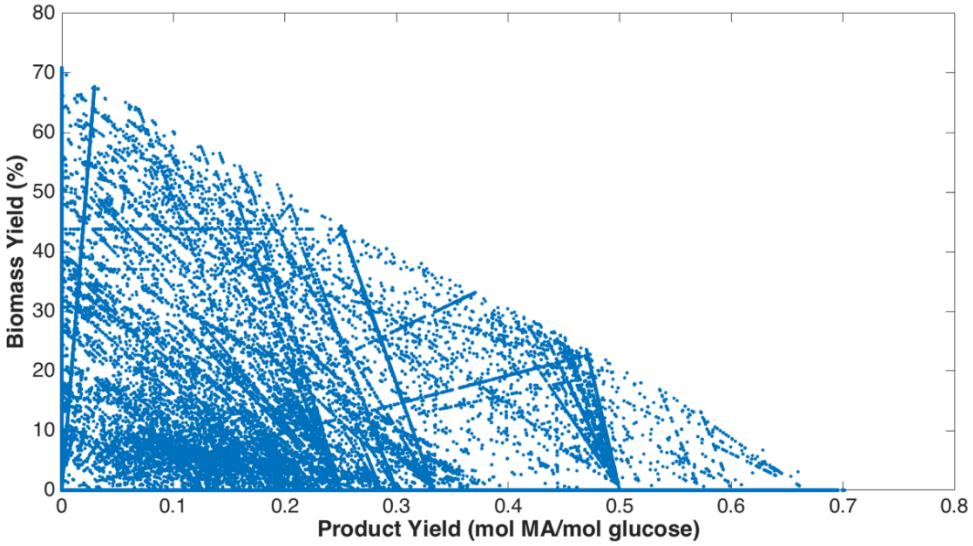


Figure B1b. Product vs. biomass yield plot of the EFM distribution of muconic acid synthesized via the pathway MA2, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

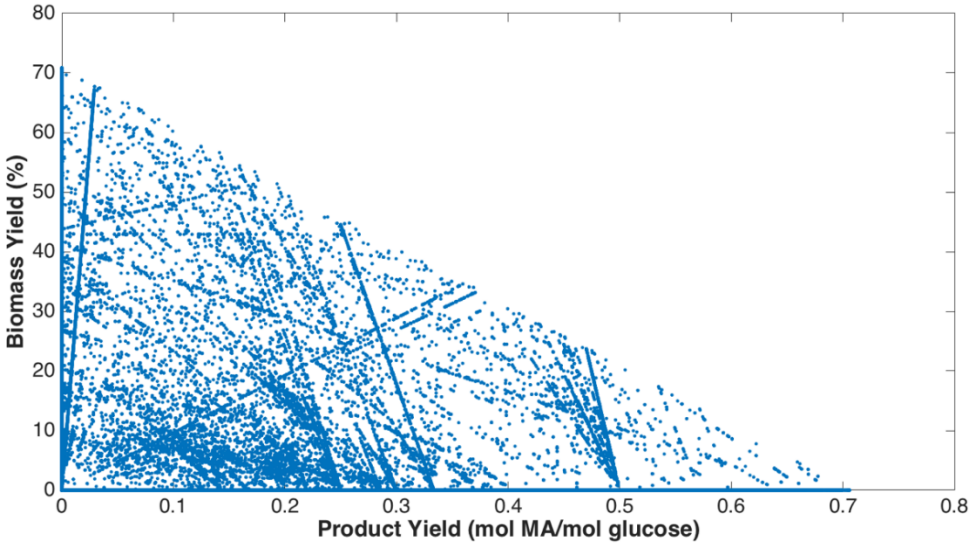


Figure B1c. Product vs. biomass yield plot of the EFM distribution of muconic acid synthesized via the pathway MA3, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

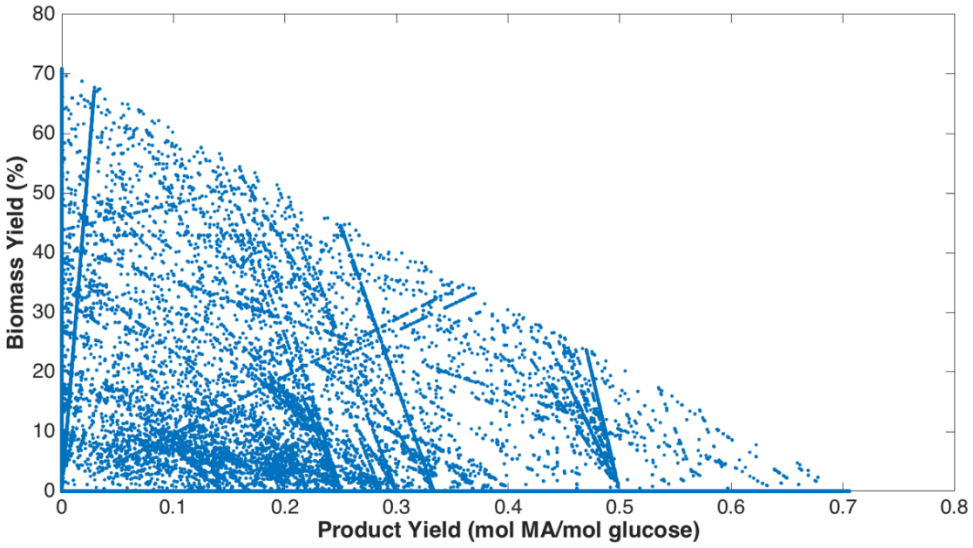


Figure B1d. Product vs. biomass yield plot of the EFM distribution of muconic acid synthesized via the pathway MA4, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

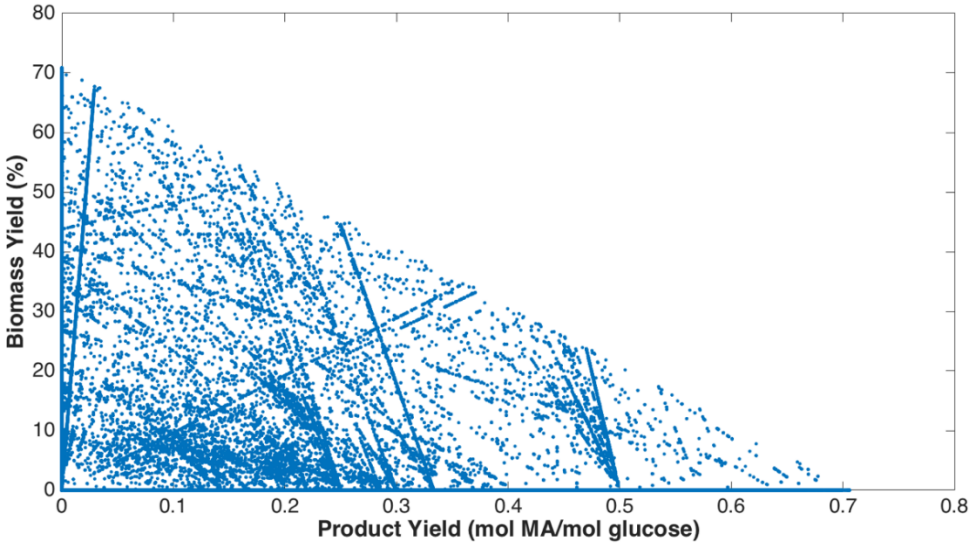


Figure B1e. Product vs. biomass yield plot of the EFM distribution of muconic acid synthesized via the pathway MA5, as predicted using the wild-type *E. coli* network. Product and biomass yields reported as the fraction of carbon in the glucose substrate converted into each product.

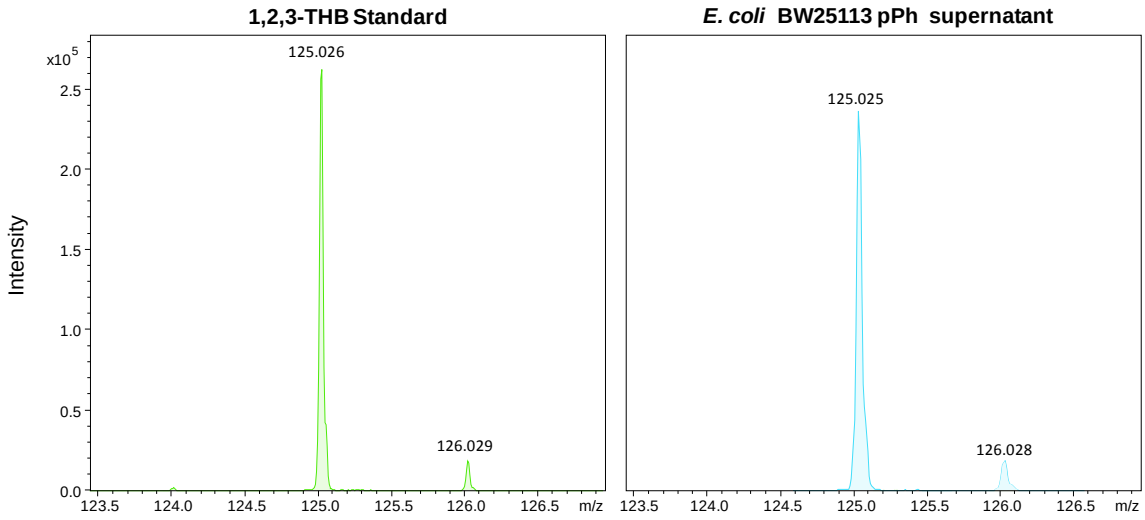


Figure B2. Detection of 1,2,3-trihydroxybenzene (1,2,3-THB) in culture supernatants from *E. coli* BW25113 whole resting cells expressing *phKLMNOP* of *Pseudomonas stutzeri* OX1. Cells were suspended in PBS buffer (pH 6.8) supplemented with 1 mM phenol.

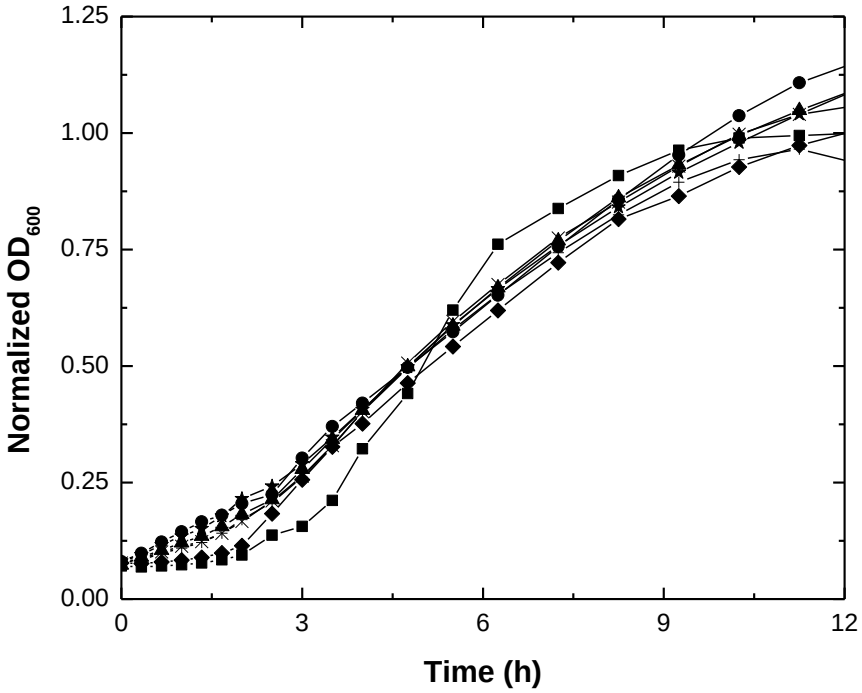


Figure B3. Effect of the exogenous addition of 1,2,3-trihydroxybenzene on the growth of *E. coli* NST74 Δ*pheA*^{fbr} when added to early exponential cultures at final concentrations of 0 g/L (■), 1 g/L (●), 2 g/L (▲), 3 g/L (♦), 4 g/L (♦), 5 g/L (x), and 6 g/L (+). Error bars represent one standard deviation from triplicate experiment.

MA1

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R10		'--> P_ex'	Phosphate in
R11		'--> O2'	Oxygen uptake
R12		'--> NH3'	Ammonia uptake
R13		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R14		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R15		'Muconate + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R16		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R17		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R18		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R19		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (EC 2.7.1.11)
R20		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R21		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R22		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R23		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R24		'ADP + 13-PG <--> ATP + 3-PG'	Phosphoglycerate kinase (EC 2.7.2.3)
R25		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R26		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R27		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R28		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R29		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R30		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R31		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R32		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R33		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R34		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R35		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R36		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R37		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R38		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R39		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R40		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R41		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (EC 2.3.3.1)
R42		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R43		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R44		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R45		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R46		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R47		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R48		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R49		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R50		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R51		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R52		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R53		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R54		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R55		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R56		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R57		'Formate --> CO2'	Formate hydrogenlyase
R58		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R59		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R60		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R61		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R62		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R63		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R64		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R65		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R66		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R67		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R68		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R69		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R70		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R71		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R72		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R73		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R74		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R75		'ATP --> ADP + P'	ATP hydrolysis
R76		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass formation and maintenance
R77		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R78		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R79		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80a		'DHS --> Protocatechuate'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81a		'Protocatechuate --> Catechol + CO2'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82a		'Catechol + O2 --> Muconate'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure B4a Networks and functions used for elementary flux mode analysis of muconic acid biosynthesis via MA1.

MA2

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R10		'--> P_ex'	Phosphate in
R11		'--> O2'	Oxygen uptake
R12		'--> NH3'	Ammonia uptake
R13		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R14		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R15		'Muconate + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R16		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R17		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R18		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R19		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R20		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R21		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R22		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R23		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R24		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R25		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R26		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R27		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R28		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R29		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R30		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R31		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R32		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R33		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R34		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R35		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R36		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R37		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R38		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R39		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R40		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R41		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R42		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R43		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R44		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R45		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R46		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R47		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R48		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R49		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R50		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R51		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R52		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R53		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R54		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R55		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R56		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R57		'Formate --> CO2'	Formate hydrogenlyase
R58		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R59		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R60		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R61		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R62		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R63		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R64		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R65		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R66		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R67		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R68		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R69		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R70		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R71		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R72		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R73		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R74		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R75		'ATP --> ADP + P'	ATP hydrolysis
R76		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R77		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R78		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R79		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80b		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81b		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82b		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83b		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84c		'Chorismate --> Prephenate'	Chorismate mutase (5.4.99.5)
R85$_{c}$		'Prephenate + NAD --> 4-Hydroxyphenylpyruvate + NADH + CO2'	Prephenate dehydrogenase (EC 1.3.1.12)
R86$_{c}$		'4-Hydroxyphenylpyruvate + GLU <--> 2-OXO + Tyrosine'	Tyrosine aminotransferase (EC 2.6.1.5)
R87$_{c}$		'Tyrosine --> PYR + Phenol'	Tyrosine pyruvate lyase
		'Phenol + NADPH + 0.5 O2 --> Catechol + NADP'	Phenol 2-monooxygenase (EC 1.14.13.7)
		'Catechol + O2 --> Muconate'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure B4b Networks and functions used for elementary flux mode analysis of muconic acid biosynthesis via MA2.

MA3

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R12		'--> P_ex'	Phosphate in
R14		'--> O2'	Oxygen uptake
R15		'--> NH3'	Ammonia uptake
R16		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R17		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R18		'Muconate + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R19		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R20		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R21		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R22		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R23		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R24		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R25		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R26		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R27		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R28		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R29		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R30		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R31		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R32		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R33		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R34		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R35		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R36		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R37		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R38		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R39		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R40		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R41		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R42		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R43		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R44		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R45		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R46		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R47		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R48		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R49		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R50		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R51		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R52		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R53		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R54		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R55		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R56		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R57		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R58		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R59		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R60		'Formate --> CO2'	Formate hydrogenlyase
R61		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R62		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R63		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R64		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R65		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R66		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R67		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R68		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R69		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R70		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R71		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R72		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R73		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R74		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R75		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R76		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R77		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R78		'ATP --> ADP + P'	ATP hydrolysis
R79		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R80		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R81		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R82		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80b		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81b		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82b		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83b		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84c		'Chorismate <--> Isochorismate'	Salicylate biosynthesis isochorismate synthase (EC 5.4.4.2)
R85c_{ij}		'Isochorismate --> Salicylate + PYR'	Isochorismate pyruvavte lyase (EC 4.2.99.21)
R86c_{ij}		'Salicylate --> Phenol + CO2'	Salicylate decarboxylase
		'Phenol + NADPH + 0.5 O2 --> Catechol + NADP'	Phenol 2-monooxygenase (EC 1.14.13.7)
		'Catechol + O2 --> Muconate'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure B4c Networks and functions used for elementary flux mode analysis of muconic acid biosynthesis via MA3.

MA4

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R12		'--> P_ex'	Phosphate in
R14		'--> O2'	Oxygen uptake
R15		'--> NH3'	Ammonia uptake
R16		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R17		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R18		'Muconate + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R19		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R20		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R21		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R22		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R23		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R24		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R25		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R26		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R27		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R28		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R29		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R30		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R31		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R32		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R33		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R34		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R35		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R36		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R37		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R38		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R39		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R40		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R41		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R42		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R43		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R44		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R45		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R46		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R47		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R48		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R49		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R50		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R51		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R52		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R53		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R54		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R55		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R56		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R57		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R58		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R59		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R60		'Formate --> CO2'	Formate hydrogenlyase
R61		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R62		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R63		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R64		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R65		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R66		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R67		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R68		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R69		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R70		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R71		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R72		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R73		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R74		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R75		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R76		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R77		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R78		'ATP --> ADP + P'	ATP hydrolysis
R79		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R80		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R81		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R82		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80b		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81b		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82b		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83b		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84d		'Chorismate --> 4-Hydroxybenzoate + PYR'	Chorismate lyase (EC 4.1.3.40)
R85d		'4-Hydroxybenzoate --> Phenol + CO2'	4-hydroxybenzoate decarboxylase
		'Phenol + NADPH + 0.5 O2 --> Catechol + NADP'	Phenol 2-monooxygenase (EC 1.14.13.7)
		'Catechol + O2 --> Muconate'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure B4d Networks and functions used for elementary flux mode analysis of muconic acid biosynthesis via MA4.

MA5

Abbreviation		Reaction	Function / Enzyme
R1	reactionFormulas = {	'--> Glucose'	Glucose feed
R2		'Product_ex -->'	Product out
R3		'Biomass_ex -->'	Biomass_ex out
R4		'CO2 -->'	Carbon dioxide out
R5		'Lactate -->'	Lactate out
R6		'Ethanol -->'	Ethanol out
R7		'Formate -->'	Formiate out
R8		'Acetate -->'	Acetate out
R9		'SUCC_ex -->'	Succinate out
R12		'--> P_ex'	Phosphate in
R14		'--> O2'	Oxygen uptake
R15		'--> NH3'	Ammonia uptake
R16		'P_ex + ATP --> ADP + 2 P'	Phosphate-transporting ATPase (EC 3.6.3.27)
R17		'3 SUCC + ATP --> 3 SUCC_ex + ADP + P'	Succinate export (ABC transporter)
R18		'Muconate + ATP --> Product_ex + ADP + P'	Product export (ABC transporter)
R19		'Glucose + PEP --> G6P + PYR'	Phosphotransferase system (EC 2.7.1.69)
R20		'Glucose + ATP --> G6P + ADP'	Hexokinase (EC 2.7.1.1)
R21		'G6P <--> F6P'	Glucose-6-phosphate isomerase (EC 5.3.1.9)
R22		'ATP + F6P --> ADP + F-16-BP'	Phosphofructokinase (2.7.1.11)
R23		'F-16-BP --> F6P + P'	Fructose 1,6-bisphosphatase (EC 3.1.3.11)
R24		'F-16-BP <--> GA3P + DHAP'	Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13)
R25		'DHAP <--> GA3P'	Triose-phosphate isomerase (EC 5.3.1.1)
R26		'GA3P + NAD + P <--> 13-PG + NADH'	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
R27		'ADP + 13-PG <--> ATP + 3-PG'	3-Phosphoglycerate phosphatase (EC 3.1.3.38)
R28		'3-PG <--> 2-PG'	Phosphoglycerate mutase (EC 5.4.2.1)
R29		'2-PG <--> PEP'	Phosphopyruvate hydratase (EC 4.2.1.11)
R30		'PEP + ADP --> PYR + ATP'	Pyruvate kinase (EC 2.7.1.40)
R31		'PYR + ATP --> PEP + AMP + P'	Phosphoenolpyruvate synthase (EC 2.7.9.2)
R32		'G6P + NADP --> GLC-LAC + NADPH'	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
R33		'GLC-LAC --> 6-P-Gluconate'	6-Phosphogluconolactonase (EC 3.1.1.31)
R34		'6-P-Gluconate --> KDPG'	Phosphogluconate dehydratase (EC 4.2.1.12)
R35		'KDPG --> PYR + GA3P'	2-Keto-3-deoxygluconate-6-phosphate aldolase (EC 4.1.2.14)
R36		'6-P-Gluconate + NADP <--> RIBU-5P + CO2 + NADPH'	Phosphogluconate dehydrogenase (EC 1.1.1.44)
R37		'RIBU-5P <--> XYL-5P'	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
R38		'RIBU-5P <--> RIBO-5P'	Ribose-5-phosphate isomerase (EC 5.3.1.6)
R39		'S7P + GA3P <--> RIBO-5P + XYL-5P'	Transketolase (EC 2.2.1.1)
R40		'S7P + GA3P <--> E-4P + F6P'	Transaldolase (EC 2.2.1.2)
R41		'F6P + GA3P <--> E-4P + XYL-5P'	Transketolase (EC 2.2.1.1)
R42		'PYR + H-CoA + NAD --> AC-CoA + NADH + CO2'	Pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, EC 1.8.1.4)
R43		'Acetate + ATP + H-CoA --> AC-CoA + AMP + 2 P'	Acetyl-CoA synthetase / Acetate-CoA ligase (EC 6.2.1.1)
R44		'AC-CoA + OAA --> CIT + H-CoA'	Citrate synthase (E.C. 2.3.3.1)
R45		'CIT <--> ICI'	Aconitase (EC 4.2.1.3)
R46		'ICI + NADP <--> 2-OXO + CO2 + NADPH'	Isocitrate dehydrogenase (NADP dependent) (EC 1.1.1.42)
R47		'2-OXO + NAD + H-CoA --> SUCC-CoA + NADH + CO2'	2-Oxoglutarate dehydrogenase complex (EC 1.2.4.2, EC 2.3.1.61, EC 1.8.1.4)
R48		'SUCC-CoA + ADP + P <--> SUCC + H-CoA + ATP'	Succinyl-CoA synthetase (EC 6.2.1.5)
R49		'SUCC + FAD --> FUM + FADH2'	Succinate dehydrogenase (EC 1.3.5.1)
R50		'FUM + FADH2 --> SUCC + FAD'	Fumarate reductase (EC 1.3.99.1)
R51		'FUM <--> MAL'	Fumarase (EC 4.2.1.2)
R52		'MAL + NAD <--> OAA + NADH'	Malate dehydrogenase (EC 1.1.1.37)
R53		'ICI --> GLYOXY + SUCC'	Isocitrate lyase (EC 4.1.3.1)
R54		'GLYOXY + AC-CoA --> MAL + H-CoA'	Malate synthase (EC 2.3.3.9)
R55		'PEP + CO2 --> OAA + P'	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)
R56		'OAA + ATP --> PEP + ADP + CO2'	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
R57		'MAL + NADP <--> PYR + CO2 + NADPH'	Malic enzyme (NADP dependent) (EC 1.1.1.40)
R58		'MAL + NAD --> PYR + CO2 + NADH'	Malic enzyme (NAD dependent) (EC 1.1.1.38)
R59		'PYR + H-CoA <--> AC-CoA + Formate'	Pyruvate formate lyase (EC 2.3.1.54)
R60		'Formate --> CO2'	Formate hydrogenlyase
R61		'PYR + NADH <--> Lactate + NAD'	Lactate dehydrogenase (EC 1.1.1.28)
R62		'AC-CoA + NADH <--> ACA + NAD + H-CoA'	Acetaldehyde dehydrogenase
R63		'ACA + NADH <--> Ethanol + NAD'	Ethanol dehydrogenase (EC 1.1.1.1)
R64		'AC-CoA + P <--> ACP + H-CoA'	Phosphate acetyltransferase (EC 2.3.1.8)
R65		'ACP + ADP <--> Acetate + ATP'	Acetate kinase (EC 3.6.1.7)
R66		'GLN + 2-OXO + NADPH --> 2 GLU + NADP'	Glutamate synthase (NADP dependent) (EC 1.4.1.13)
R67		'2-OXO + NH3 + NADPH <--> GLU + NADP'	Glutamate dehydrogenase (NADP dependent) (EC 1.4.1.4)
R68		'GLU + NH3 + ATP --> GLN + ADP + P'	Glutamine synthetase (EC 6.3.1.2)
R69		'GLN --> GLU + NH3'	Glutaminase (EC 3.5.1.2)
R70		'NADPH + NAD --> NADP + NADH'	NAD(P)+ transhydrogenase (EC 1.6.1.1)
R71		'NADH + FAD --> NAD + FADH2'	Flavin reductase (NAD dependent) (EC 1.5.1.36)
R72		'NADP + NADH + H[e] --> NADPH + NAD + H[c]'	NAD(P)+ transhydrogenase (EC 1.6.1.2)
R73		'NADH + UQ8 + 4 H[c] --> NAD + UQ8H2 + 4 H[e]'	NADH dehydrogenase
R74		'FADH2 + UQ8 --> FAD + UQ8H2'	FADH2 dehydrogenase
R75		'UQ8H2 + 4 H[c] + 0.5 O2 --> UQ8 + 4 H[e]'	Cyt_b0 / Cyt_bd oxidase
R76		'ADP + 4 H[e] + P <--> ATP + 4 H[c]'	ATP synthase
R77		'AMP + ATP <--> 2 ADP'	Adenylate kinase (EC 2.7.4.3)
R78		'ATP --> ADP + P'	ATP hydrolysis
R79		'6965 NH3 + 206 G6P + 72 F6P + 627 RIBO-5P + 361 E-4P + 129 GA3P + 1338 3-PG + 720 PEP + 2861 PYR + 2930 AC-CoA + 1481 OAA + 1078 2-OXO + 16548 NADPH + 56357 ATP + 3548 NAD --> Biomass_ex + 16548 NADP + 2930 H-CoA + 1678 CO2 + 56357 ADP + 56357 P + 3548 NADH'	Biomass_ex formation and maintenance
R80		'E-4P + PEP --> DAHP + P'	3-deoxy-7-Phosphoheptulonate synthase (EC 2.5.1.54)
R81		'DAHP --> DHQ + P'	3-Dehydroquinate synthase (EC 4.2.3.4)
R82		'DHQ --> DHS'	3-Dehydroquinate dehydratase (EC 4.2.1.10)
R80b		'DHS + NADPH --> Shikimate + NADP'	3-Dehydroshikimate dehydratase (EC 4.2.1.118)
R81b		'Shikimate + ATP --> Shikimate-3-P + ADP'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82b		'Shikimate-3-P + PEP --> Carboxyvinyl-Shikimate-3-P + P'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
R83b		'Carboxyvinyl-Shikimate-3-P --> Chorismate + P'	Chorismate synthase (EC 4.2.3.5)
R84d		'Chorismate --> 4-Hydroxybenzoate + PYR'	Chorismate lyase (EC 4.1.3.40)
R85d		'4-Hydroxybenzoate + NADPH --> Protocatechuate + NADP'	4-hydroxybenzoate 3-monooxygenase (EC 1.14.13.2/EC 1.14.13.33)
R81a		'Protocatechuate --> Catechol + CO2'	Protocatechuate decarboxylase (EC 4.1.1.63)
R82a		'Catechol + O2 --> Muconate'	Catechol 1,2-dioxygenase (EC 1.13.11.1)
	};		
	mnet = CalculateFluxModes(reactionFormulas);		

Figure B4e Networks and functions used for elementary flux mode analysis of muconic acid biosynthesis via MA5.

APPENDIX C

SUPPLEMENTAL INFORMATION TO EXPLORING A SYNTHETIC METABOLIC FUNNELING APPROACH TO ENHANCE PHENOL AND MUCONIC ACID BIOSYNTHESIS

Table C1 List of primers designed and used for phenol and muconic acid production

Primer	Sequence (5' → 3')
qsuB -- BamHI - F	ATA GGA TCC AGG AGG ATA AAT AAT GCG TAC ATC CAT TGC CAC TGT TTG
qsuB -- HindIII - R	ATT AAG CTT CTA GTT TGG GAT TCC CCG CTC GA
catA -- BglII-F	ATA AGA TCT AGG AGG ATA AAT AAT GAC CGT GAA AAT TTC CCA C
catA -- XbaI - R	ATT TCT AGA TCA GCC CTC CTG CAA CGC
ubiC -- EcoRI - F	ATA GAA TTC AGG AGG ATA AAT AAT GTC ACA CCC CGC GTT AAC G
ubiC -- BglII - R	ATT AGA TCT TTA GTA CAA CGG TGA CGC CGG TAA A
kpdBCD -- BamHI - F	ATA GGA TCC CCC GTC CGG AGA GGG TAA TTT AAA TAT AAA GTT CG
kpdBCD -- HindIII - R	ATT AAG CTT CTT AGC GGG CCC CTT TAT TAA CGC T
pY3 [tutA] - F	CTG CAC GCT TTG ACT ATA TCT AAG GAT CCA AAC TCG AGT AAG G
pY3 [tyrA] - R	CAT GGA TGG CCT CCT AGA TCT TTT GAA TTC TGA AAT TGT TAT C
tyrA [pY3] - F	GAT AAC AAT TTC AGA ATT CAA AAG ATC TAG GAG GCC ATC CAT G
tyrA [tutA] - R	CGG ATA ATT CAT TAT TTA TCC TCC TTT AGA TCC TTA CTG GCG ATT
tutA [tyrA] - F	AAT CGC CAG TAA GGA TCT AAA GGA GGA TAA ATA ATG AAT TAT CCG
tutA [pY3] - R	CCT TAC TCG AGT TTG GAT CCT TAG ATA TAG TCA AAG CGT GCA G
SDC-pchB-entC [pS3] -- F	AAG GAG GCC ATC CAT GCG TGG TAA AGT TAG C
SDC-pchB-entC [pS3] -- R	GTT TGG ATC CTT AAT GCA ATC CAA AAA CG
pS3 [SDC-pchB-entC] -- F	ATT GCA TTA AGG ATC CAA ACT CGA GTA AG
pS3 [SDC-pchB-entC] -- R	CTT TAC CAC GCA TGG ATG GCC TCC TTT AGA TC
pheA (Keio) -- F	CGT GTG AAA CAG AAT GCG AAG ACG AAC AAT A
pheA (Keio) -- R	TAA TCC AGT GCC GGA TGA TTC ACA TCA TC
pykA (Keio) -- F	ATC GCG GCG TTA TTT CAT TCG GAT T
pykA (Keio) -- R	AAC TGT AGG CCG GAT GTG GC
pykF (Keio) -- F	GCG AGG CAC CAC CAC TTT CG
pykF (Keio) -- R	AGC GCC CAT CAG GGC G
crr (Keio) -- F	CTA TGA GCG CCA TTT CTA TCC CGC GC
crr (Keio) -- R	CCT GAA AGG GAC TGG CGA CCT G

APPENDIX D

PRELIMINARY DATA SUPPORTING THE ENGINEERING OF A DYNAMIC PHENOTYPE CONTROL SWITCH

Investigating Transcriptional Control Using a Metabolic Toggle Switch

To improve MA biosynthesis, a metabolic toggle switch was constructed (Figure D1 and Figure D2) to modulate expression between native shikimate dehydrogenase (encoded by *aroE*) and the original heterologous muconic acid biosynthesis pathway (encoded by *qsuB*, *aroY*, and *catA*). As illustrated in Figure D2, the toggle will express the MA pathway and repress expression of *aroE* in the presence of aTc. In contrast, the toggle will express *aroE* and repress MA pathway expression in the presence of IPTG. To test functionality, said toggle was transformed into *E. coli* NST74 Δ*aroE* and evaluated against appropriate control strains (Table D1). Strains were cultured at 32ºC in baffled shake flasks using 50 mL of a phosphate-limited minimal media (MM1) composed of (concentrations in parentheses, all in g/L): $MgSO_4 \cdot 7H_2O$ (0.5), $(NH_4)_2SO_4$ (4.0), MOPS (24.7), KH_2PO_4 (0.3), K_2HPO_4 (0.7), $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ ($3.7 \cdot 10^{-4}$), H_3BO_3 ($2.5 \cdot 10^{-3}$), $CoCl_2 \cdot 6H_2O$ ($7.14 \cdot 10^{-4}$), $CuSO_4$ ($1.6 \cdot 10^{-4}$), $MnCl_2 \cdot 4H_2O$ ($1.6 \cdot 10^{-3}$), $ZnSO_4 \cdot 7H_2O$ ($2.88 \cdot 10^{-4}$), $FeCl_3$ ($5.0 \cdot 10^{-5}$), and glucose (20), supplemented with appropriate antibiotics. *E. coli* NST74 Δ*aroE* strains lacking the metabolic toggle plasmid were additionally supplemented with 100 mg/L each of Phe, Tyr, Trp, and 50 mg/L each of *p*-hydroxybenzoate, *p*-aminobenzoate, and 2,3-dihydroxybenzoate. To induce *aroE* or *qsuB-aroY-catA* expression, IPTG and/or aTc were added at final concentrations of 0.4 mM and 400 nM, respectively, when cells reached an OD_{600} of ~ 0.7. Although MA was produced using the metabolic toggle design depicted in Figure D1, significant room for improvement remains. Additional characterization should be investigated initially focused on GFP-mCherry levels, varying levels of inducer, and varying time point at which inducer is added.

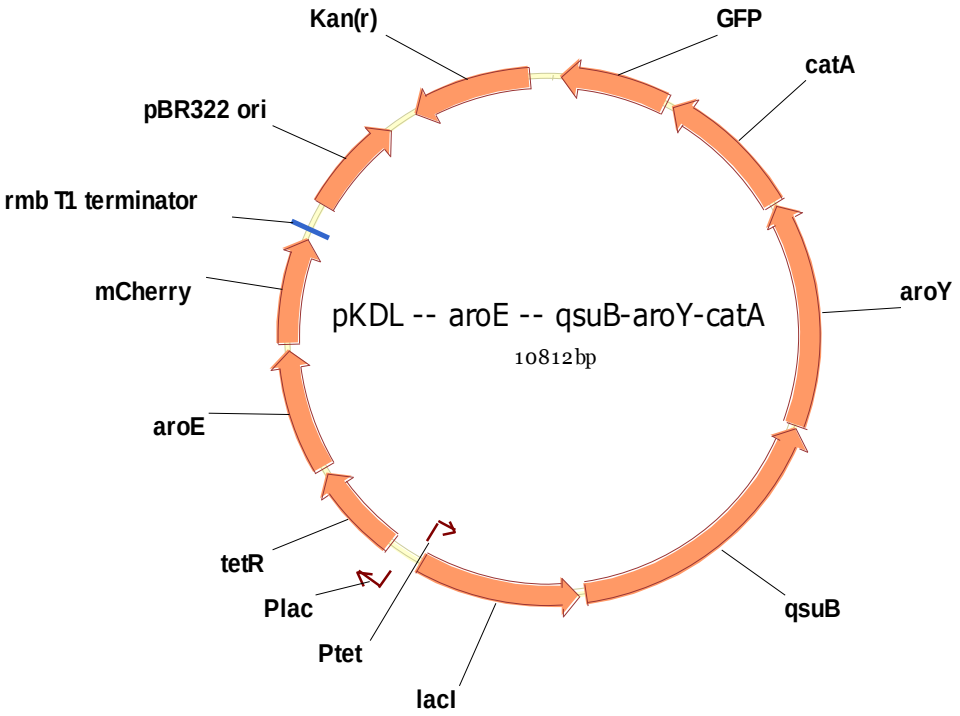


Figure D1 Plasmid map illustrating the metabolic toggle switch for MA production

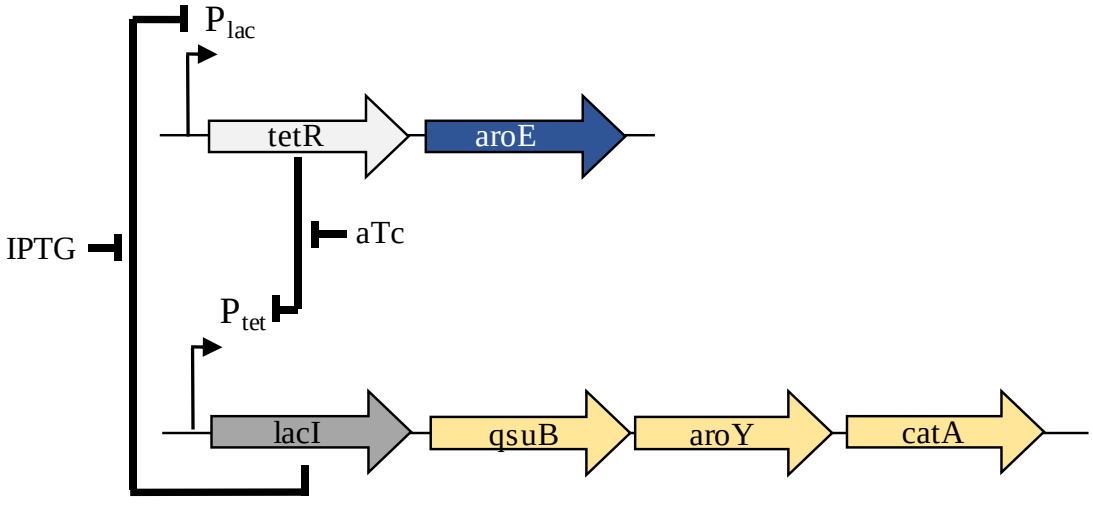


Figure D2 Engineered metabolic toggle switch design. Addition of aTc results in MA pathway (*qsuB-aroY-catA*) expression, while addition of IPTG results in native *aroE* expression.

Table D1 Plasmids used in this study

Plasmid	Description	Source
pKDL	pBR322 ori, Kan^r, P_{tet}-lacI-GFP, P_{lac}-tetR-mCherry	Xiao Wang Lab (ASU)
pMA	pBR322 ori, Kan^r, P_{tet}-*lacI-qsuB-aroY-catA*-GFP, P_{lac}-*tetR-aroE*-mCherry	This study

Table D2 MA production using metabolic toggle switch

Strain	Plasmid	Inducer Added	OD_{600}	MA (mg/L)
NST74 Δ*aroE*	--	none	3.2	n.d.
NST74 Δ*aroE*	pMA	none	1.2	n.d.
NST74 Δ*aroE*	pMA	aTc	2.2	229
NST74 Δ*aroE*	pMA	IPTG	2.1	122

Investigating Transcriptional Control Using CRISPRi

To improve MA biosynthesis, a CRISPRi plasmid combination was engineered to transcriptionally repress native shikimate dehydrogenase (encoded by *aroE*). To achieve repression, one plasmid harbors the custom designed sgRNA sequence (ATG GAA ACC TAT GCT GTT TT) under control of the strong constitutive promoter J23119 (Figure D3) while *dcas9* expression was under control of the inducible P_{tet} promoter (Figure D4). In addition, MA production was achieved via expression of *qsuB-aroY-catA* under control of the inducible P_{trc} promoter (Figure D5). To test functionality, the two CRISPRi vectors and the MA pathway vector were transformed into *E. coli* NST74 and evaluated against appropriate control strains (Table D3). Strains were cultured at 32ºC in baffled shake flasks using 50 mL of a phosphate-limited minimal media (MM1) composed of (concentrations in parentheses, all in g/L): $MgSO_4 \cdot 7H_2O$ (0.5), $(NH_4)_2SO_4$ (4.0), MOPS (24.7), KH_2PO_4 (0.3), K_2HPO_4 (0.7), $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ ($3.7 \cdot 10^{-4}$), H_3BO_3 ($2.5 \cdot 10^{-3}$), $CoCl_2 \cdot 6H_2O$ ($7.14 \cdot 10^{-4}$), $CuSO_4$ ($1.6 \cdot 10^{-4}$), $MnCl_2 \cdot 4H_2O$ ($1.6 \cdot 10^{-3}$), $ZnSO_4 \cdot 7H_2O$ ($2.88 \cdot 10^{-4}$), $FeCl_3$ ($5.0 \cdot 10^{-5}$), and glucose (20), supplemented with appropriate antibiotics. To induce CRISPRi-mediated *aroE* repression and *qsuB-aroY-*

catA expression, aTc and/or IPTG were added at final concentrations of 400 nM and 0.4 mM, respectively, when cells reached an OD_{600} of ~ 0.7. Although MA production remained unchanged with the use of CRISPRi, significant room for improvement remains. Additional characterization should be investigated initially focused varying induction timing and alternative expression platforms for the sgRNA and d*cas9* components.

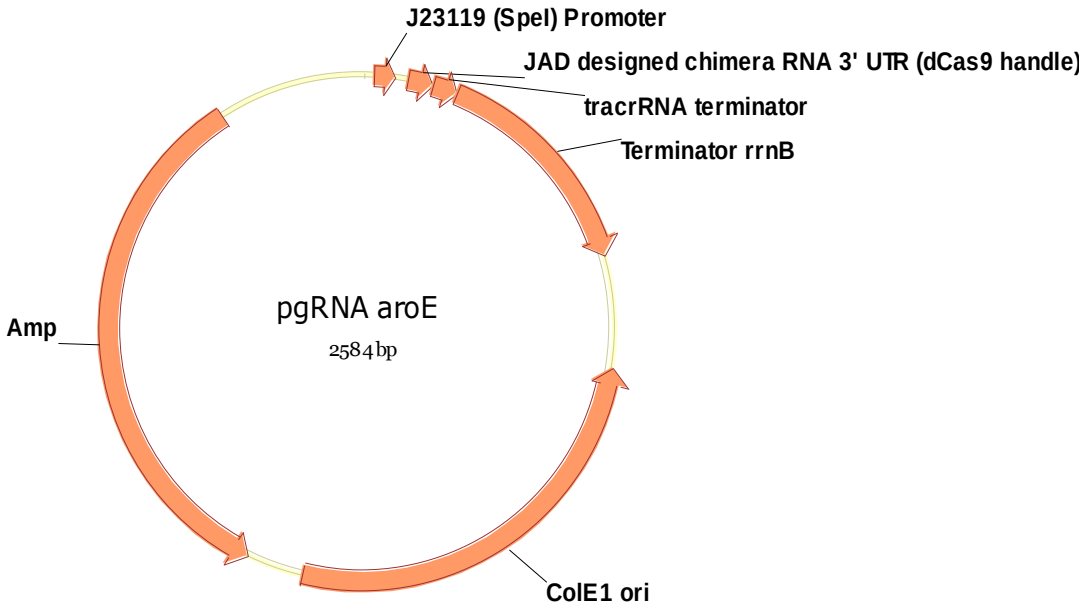


Figure D3 Plasmid map for sgRNA expression targeting native *aroE*.

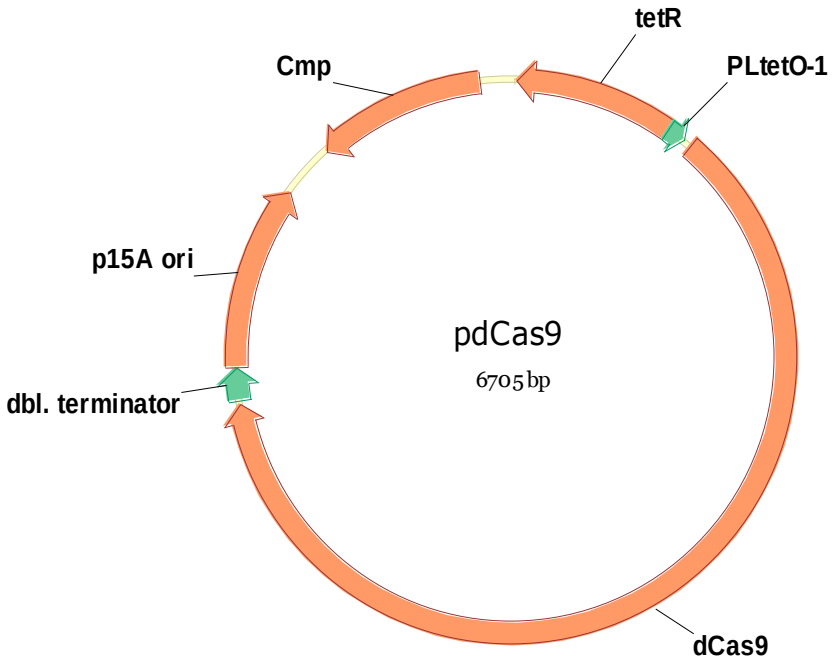


Figure D4 Plasmid map for *dcas9* expression.

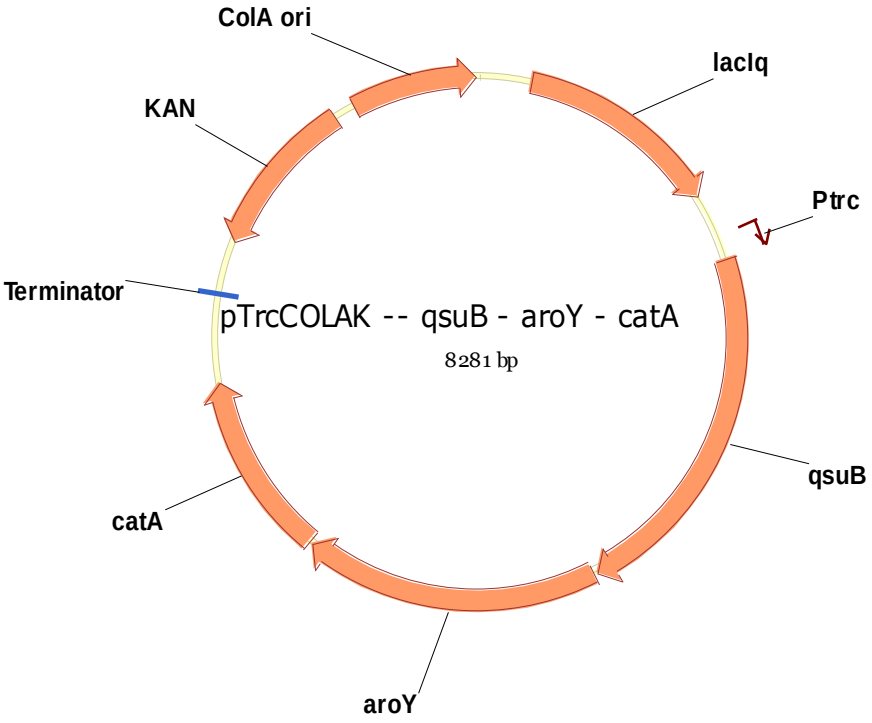


Figure D5 Plasmid map for MA production via expression of *qsuB-aroY-catA*.

Table D3 Plasmids used in this study

Plasmid	Description	Source
pgRNA	colE1 ori, Amp^r, P_{J23119}-gRNA(aroE)	This study
pdCAS9	p15A ori, Cmp^r, tetR-P_{tet}-dCAS9	This study
pMA	ColA ori, Kan^r, lacIq-P_{trc}-*qsuB-aroY-catA*	This study

Table D4 MA production using CRISPRi

Strain	Plasmid	OD_{600}	Phenylalanine (mg/L)	MA (mg/L)
NST74	--	3.0	1150	n.d.
NST74	pMA	2.7	539	1592
NST74	pgRNA, pdCAS9	2.8	796	n.d.
NST74	pgRNA, pdCAS9, pMA	2.5	458	1452